AD 285 065

Errata Sheet for Research Memorandum 45

in section: "The Validity of Economic Growth Statistics and Growth Rates"

Page 8, line 2 - change "high" to "low"

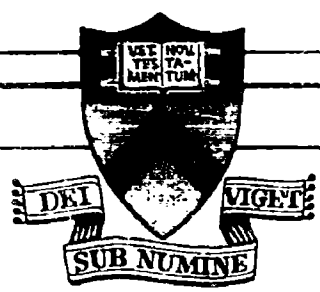
Page 8, line 3 - change "low" to "high"

Page 9, the second factor of the numerator of the fraction under the nth root sign, used in the estimation of the rate of growth should read:

$$(1 + \frac{r}{100})^n$$

rather than:

$$(1 + \frac{r}{100})$$



PRINCETON UNIVERSITY

Issued Under Office of Naval Research
Contract No. Nonr- 1858(16)

ON THE ACCURACY OF NATIONAL INCOME AND GROWTH STATISTICS

Oskar Morgenstern

Econometric Research Program
Research Memorandum No. 43
1 August 1962

The research described in this paper was supported by the Office of Naval Research. Reproduction, translation, publication, use and disposal in whole or in part by or for the United States Government is permitted.

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N. J.

PREFACE

The following two papers are consecutive chapters from the completely revised and rewritten second edition of my book, On the Accuracy of Economic Observations. The first edition of this work appeared at the Princeton University Press in 1950, and has been out of print since 1952. The manuscript of the revised edition will be in the printer's hands within a few weeks.

It gives me great pleasure to thank my collaborators in the Econometric Research Program, especially Nevins D. Baxter, John G. Cragg, Morton D. Davis, and Dorothy D. Green, for valuable assistance rendered in the preparation of this material. Lois A. Crooks was responsible for the typing, which she has done in the same competent manner as in all other of our current research memoranda.

Oskar Morgenstern

July 9, 1962.

ON THE ACCURACY OF

NATIONAL INCOME STATISTICS

Contents

1. Introduction.

> The official national income estimates prepared by the ... Department of Commerce ... have become widely accepted in this country and abroad by professional and lay users alike. The publication of new estimates is front page news ... and the business and financial world eagerly awaits their arrival. Such a reception of a body of economic data is almost unprecedented — it can be explained only on the basis that income and product accounts help satisfy the urgent need for informative statistical data on current economic developments ... . The wide use of the accounts places a serious obligation on national income analysts both inside and outside the government. The concepts, methodology and estimates underlying the estimates affect their meaning and their movements, and deficiencies in any one of these elements can lead to misleading results.[1]

In this section we shall examine the question of the accuracy of the national income estimates. Many of our comments also apply to aggregative measures in general, although we shall be specifically dealing with the difficulties confronted in the estimation of national income.

Our concern will be primarily with data for the United States, because they are plentiful; and American writers have been pioneers in the establishment of national income statistics. The problems can be most clearly seen in examining American statistics. In a later section (8), we shall also refer to British data and their errors; there the situation is similar: great efforts, important contributions, eminent authors involved. Yet there are like results: great errors, many revisions of estimates, lack of convergence of the revisions. Space forbids going into the study of the statistics of additional countries. But it is safe to say that it would be only by extraordinary circumstances that the efforts in other countries were more successful. In fact, it is clear that if the United States and the United Kingdom cannot produce better national income statistics than they do, those of others will, in general, not even come near

---

[1] A Critique of the United States Income and Product Accounts, Studies in Income and Wealth, Vol. 22, National Bureau of Economic Research, Princeton, 1958, p. 3.

their quality. This is certainly true of the underdeveloped regions of the world, which, though not containing the largest parts of the world's incomes, do comprise the greatest number of people. It is simply technically impossible for national income statistics of Africa, South and Central America, and Asia to be better than those of the United States and the United Kingdom. Even if a few were in some specific sense of the same quality, that would neither satisfy nor affect the specific problems of their comparability (see below, 9).

In recent years, the trend in economics has been toward the collection of vast aggregates of data. These aggregates present large problems in estimation,[1] for they require data on parts of the economy which are not fully explored and about which there is still little precise information. Therefore all sorts of ingenious (and often inexact) devices are employed for arriving at estimates of these relatively unknown components. As a result of these uncertainties, the aggregates are frequently of dubious accuracy. All this is mitigated only by a general hope — seldom specified and never proved — that errors in components will cancel out and that we would get an acceptable total. Whether the errors are "large" or not depends, of course, on the uses to which the statistics are to be put, as will be mentioned more specifically below.

---

[1] In view of the often immense practical-political consequences of showing one rather than another figure for national income, in order to arrive at a "suitable" number the use of deceptive and political methods similar to those mentioned in Part I is not unheard of. So the national income of Japan was negotiated between the Japanese Government and the American Occupation Forces shortly after the last war, as reported by M. Bronfenbrenner in his review of the first edition of the present book (cf. Land Economics, February 1952, p. 82). The reason was, of course, that the amount agreed upon influenced the size of economic assistance by the United States. Have these figures entered later econometric research? What other, similar, instances would be mentioned?

2. Concepts of National Income.

In the notion of a "national income" most difficulties of economics culminate. The "Wealth of Nations" has been the prime concern of economists as long as there has been any systematic writing in economics, and so it will be for the future. Neither the conceptual nor the statistical problems in this field have been resolved to anyone's satisfaction, though a great deal of progress has been made in both respects. The two areas are interdependent, since nothing can be measured for which there exist no good concepts, and concepts, no matter how precise, are of little practical value if the corresponding measurements cannot be performed. The literature in this field is immense and comprises most of the famous names of the discipline. It is clearly impossible to survey it, even to list the most important works.[1] We shall limit ourselves to a brief discussion of those principal points that are of direct relevance for the evaluation of errors of measurement and reliability of basic data. Besides S. Kuznets, at least the names of M. A. Copeland, M. Gilbert, G. Jaszi, and I Kravis should be mentioned for this country. All have in various ways beneficially influenced the National Income Division of the Department of Commerce, the basic source of all American figures.

[1] Apart from early classical works and the literature associated with I. Fisher, A. C. Pigou, V. Pareto and others, there are the efforts of the National Bureau of Economic Research, where S. Kuznets has been — and still is — one of the prime movers. A survey, still of value, is Kuznets' "National Income," Encyclopedia of the Social Sciences (1933), reprinted in Readings in the Theory of Income Distribution (American Economic Association, 1946), where there is found an extensive Bibliography. Particular mention is deserved by the important and massive work of Paul Studenski, The Income of Nations, Theory, Measurement, and Analysis: Past and Present, New York, 1958. The book contains vast amounts of valuable information about the history, methods and results of income measurements. The question of accuracy for different methods and countries is discussed on pp. 254-264. For the rest we refer to the more than 25 volumes, Studies in Income and Wealth, (National Bureau of Economic Research). This large number speaks tellingly of the effort that has been made in the United States alone. Similar efforts have been made in Great Britain and other countries, where gradually more means are being put to the task as the economies expand, new scholarly talent becomes available, and the uses increase to which governments and private interests try to put these statistics.

In the following no attempt is made even to list all conceptual difficulties. Instead only those are mentioned where problems of errors arise strongly. Some of these problems are discussed elsewhere in this study (e.g., regarding the volume and value of agricultural or mining output, or regarding the reliability of price statistics, etc.), but even for these our mention is only of an illustrative nature.

We can limit ourselves to the few following comments on concepts because we assume the reader to be reasonably well acquainted with the manner in which the gross national product is obtained, and how it differs from net national product, national income, personal income, etc. But in order to have the respective relationships and magnitudes before us, Table N-1 shows the composition of gross national product for 1960.

(a) Imputation of value, perhaps the classical problem of economic theory, had not been resolved until in the theory of games a satisfactory solution was found.[1] In the present case the problem is to assign to goods and services produced a measure largely derived from money flows. But if money does not flow, this does not mean that income which is not being recorded is not being generated. This is an old problem. A classical illustration is that of persons living in houses they own themselves. If these same houses were owned by others rent would have to be paid (in money, goods, or services), thereby swelling the national product. To avoid this, a value has to be imputed to owner-occupancy. This is, obviously, a tricky affair, with less certain results than finding out about rent payments made in money. These estimates are uncertain and many arbitrary decisions have to be made. To speak here of "accuracy" is difficult, since alternative records and procedures do not exist; the question is primarily one of

[1] By means of the characteristic function of an n-person game, expressing the non-additivity of the phenomenon of value. From there it is, of course, still a long way to the statistics of national income!

Table N-1

Composition of Gross National Product
1961
($ billion)

Gross national product*		518.7
Less:	Capital consumption allowances	45.3
Equals:	Net national product	473.4
Less:	Indirect business tax and nontax liability	48.2
	Business transfer payments	2.1
	Statistical discrepancy	- 3.1
Plus:	Subsidies minus current surplus of government enterprise	1.7
Equals:	National income	427.8
Less:	Corporate profits and inventory valuation adjustment	45.5
	Contributions for social insurance	21.6
	Excess of wage accruals over disbursements	0
Plus:	Government transfer payments to persons	31.3
	Net interest paid by government	7.3
	Dividends	15.0
	Business transfer payments	2.1
Equals:	Personal income	416.4
Less:	Personal tax and nontax payments	52.8
Equals:	Disposable personal income	363.6
Less:	Personal consumption expenditures	338.1
Equals:	Personal savings	25.6

Source: Survey of Current Business, July 1962, pp. 6-8.

*Individual items may not add to totals due to rounding.

procedure. Other items have to be treated similarly, e.g., not marketed food consumed on farms. This item should show a sharp decrease (a) with a falling farm population and (b) with agricultural production geared for sale and farmers buying larger parts of their food in stores. At earlier times (and now still in other countries) this would have been quite different. To the extent that the National Income Division uses a limited welfare concept, i.e., imputes only for few such cases, it depresses the total. The more price and market-oriented a country is, the smaller is the imputation problem; on the other hand, in the United States most people live in owner occupied houses for which "rent" is a very uncertain figure. The less developed a country is, the larger looms the imputation problem. In those (underdeveloped) countries statistics in general are of poorer quality, thus compounding the problem.

Difficulties of this type are quite common; they become especially important when comparisons over time or among differently organized countries are to be made. Trivial as it may seem, the disappearance of domestic help increases the imputation problem. Insofar as domestic help is replaced by housewives' labor which does not involve money payments, national income will be depressed. Of course, if the domestic help released enters the industrial labor force, national income will rise again, perhaps even above its former level. Obviously, there will be errors in both counts, and their nature and distribution will remain obscure. There are countless other illustrations.

(b) The treatment of government services involves another characteristic difficulty. National income is a monetary magnitude; for the private sector, the pricing mechanism is a suitable mechanism by which to value the final amount of goods and services. For the government sector, the pricing mechanism often offers no clue about the value of final output. For example,

how should the contribution to the national product made by expenditures on national defense, on highways, or on schools, etc., be evaluated? A related problem is the question of the "intermediate output" of the government. It is sometimes argued that part of the government output should not be treated as final, since it has utility only insofar as it affects the private sector — e.g., government highway building derives much of its value from the use that business makes of roads, say in the distribution of goods. It forms the overhead of society's capital. If this were somehow valued, there would be the additional problem of depreciation of this capital investment. These problems are familiar from Capital Theory and the Theory of Public Finance where the role of "social" capital has always been one still awaiting resolution.

Defense expenditures pose another difficulty because they raise the question: Is there any positive utility gained from national defense? Those who accept the welfare concept of the treatment of government in the national accounts believe that national defense does not add to the standard of living. Such expenditures would therefore merely represent a kind of intermediate expenditure, necessary to provide the conditions in which the private sector can adequately function. Under the welfare concept, therefore, the output of the governmental sector should not be valued at factor cost as is the current practice; rather that proportion which does not go toward increasing the standard of living should be considered as intermediate and therefore not be included. In general, the National Income Division totally rejects this viewpoint and argues that government output should best be valued at factor cost. Kuznets, on the other hand, is much closer to "welfare concepts" and therefore believes that some provision for double counting should be made.[1]

[1] S. Kuznets, "Discussion of the New Department of Commerce Income

(c) The main problem in getting from gross national product to national income is posed by the depreciation allowances. (The other element, indirect taxes, is determined with relative accuracy.) Estimates of depreciation are made by corporations themselves, guided by the rather unrealistic assumptions underlying the tax laws and their own often inappropriate ideas (e.g., lack of understanding of the process of inflations). It would be difficult to argue that such methods as "double declining balance" and "sum of the years digits," lifo and fifo, present a realistic appraisal of the depreciation of capital that actually takes place in the economy in a given time interval. The effects of price changes on profits and depreciation estimates are a further problem.

Another main conceptual difficulty which leads to difficulties in making the actual estimates is the valuation of services performed by financial intermediaries and the imputed interest that arises therefrom.

This list could be lengthened greatly. What has been said suffices, however, to show that conceptual differences held among statisticians at different times and in different countries are bound to have decisive influence upon these statistics. Depending on the choice of one concept rather than of another, the phenomena thus defined have their own error characteristics.

Given these difficulties, it is easy to understand that conceptual changes are frequent. They account for many of the almost continuous revisions, some of them very substantial, as inspection of Table N-2 will show. One must also ask whether it is the constantly changing nature of the economy that calls for these conceptual revisions or whether they are an expression of our inability to settle conceptual issues.

Series, National Income: A New Version," Review of Economics and Statistics XXX, August 1948, pp. 151-179. See also Volume 22 of the National Bureau of Economic Research Series, Studies in Income and Wealth, for an explanation and justification of Kuznets' definition of government output as tax revenues minus transfers.

Table N-2

U. S. National Income Revisions
($ billions)

(As reported in Survey of Current Business)

	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961
Feb. '48	202.6														
July '48	*202.5														
Feb. '49	202.5	224.4													
July '49	*201.7	*226.2													
Feb. '50	201.7	226.2	221.5												
July '50	*198.7	*223.5	*216.8												
Feb. '51	198.7	223.5	216.8	235.6											
July '51	198.7	223.5	*216.7	*239.0											
**1951	198.7	223.5	216.7	239.0											
Feb. '52	198.7	223.5	216.7	239.0	275.8										
July '52	198.7	223.5	*216.3	*239.2	*277.6										
Feb. '53	198.7	223.5	216.3	239.2	277.6	290.4									
July '53	198.7	223.5	216.3	*240.6	*278.4	*291.6									
Feb. '54	198.7	223.5	216.3	240.6	278.4	291.6	307.7								
July '54	*197.2	*221.6	*216.2	*240.0	*277.0	*291.0	*305.0								
Feb. '55	197.2	221.6	216.2	240.0	277.0	291.0	305.0	300.0							
July '55	197.2	221.6	216.2	240.0	277.0	*289.6	*303.6	*299.7							
Feb. '56	197.2	221.6	216.2	240.0	277.0	*289.5	303.6	299.7	322.3						
July '56	197.2	221.6	216.2	240.0	277.0	*290.2	*302.1	*298.3	*324.0						
Feb. '57	197.2	221.6	216.2	240.0	277.0	290.2	302.1	298.3	324.0	342.4					
July '57	197.2	221.6	216.2	240.0	277.0	290.2	302.1	*299.0	*324.1	*343.6					
Feb. '58	197.2	221.6	216.2	240.0	277.0	290.2	302.1	299.0	324.1	343.6	358.0				
***Jul. '58	*198.2	*223.5	*217.7	*241.9	*279.3	*292.2	*305.6	*301.8	*330.2	*349.4	*364.0				
Feb. '59	198.2	223.5	217.7	241.9	279.3	292.2	305.6	301.8	330.2	349.4	364.0	360.5			
July '59	198.2	223.5	217.7	241.9	279.3	292.2	305.6	301.8	330.2	*350.8	*366.5	*366.2			
Feb. '60	198.2	223.5	217.7	241.9	279.3	292.2	305.6	301.8	330.2	350.8	366.5	366.2	398.5		
July '60	198.2	223.5	217.7	241.9	279.3	292.2	305.6	301.8	330.2	350.8	*366.9	*367.7	*399.6		
Feb. '61	198.2	223.5	217.7	241.9	279.3	292.2	305.6	301.8	330.2	350.8	366.9	367.7	399.6		
July '61	198.2	223.5	217.7	241.9	279.3	292.2	305.6	301.8	330.2	350.8	366.9	*367.4	399.6	417.1	
Apr. '62	198.2	223.5	217.7	241.9	279.3	292.2	305.6	301.8	330.2	350.8	366.9	367.4	399.6	417.1	430.2

* Indicates change
** 1951 Supplement to Survey of Current Business
*** Major revision in 1958.

3. Types of Errors.

There are three principal types of error in the statistics of national income. First, there are the errors introduced in the basic data of production or expenditure for the separate industries and other economic activities. These data may arise from sampling investigations — in which case there would be the usual statistical sampling errors — or from mass enumeration. There will be difficulties in taking the proper count. Studies of the accuracy of foreign trade, mining, and agriculture give an idea of the substantial magnitude of the errors to be expected in these components.

Second, error may be produced independently of enumeration or sampling difficulties. These errors result from the effort to fit the available statistics to the conceptual framework of the aggregate. The accuracy and the success of an estimate is conditioned by the quality and quantity of the primary data. In some cases the existing data are not collected in a form directly suitable for use in estimating gross national product or one of its component items. For example, the Census Bureau's industrial enumerations and sample surveys do not provide adequate information on industry purchases of intermediate goods which must be netted from industry sales figures for gross national product purposes.[1] Most of the national income and product estimates are based on government-produced statistics, which must be assembled and adjusted to build up the income and output measures. For example, some data that become available from government agencies are a byproduct of their administrative functions.[2] This is frequently a very strong reason to suspect the quality of data obtained in this manner.

---

[1] U. S. Department of Commerce, Office of Business Economics, U. S. Income and Output, A Supplement to the Survey of Current Business, Washington, 1958, p. 66. (This will hereafter be cited as U. S. Income and Output.)

[2] U. S. Income and Output, p. 70.

Third, since not all basic data are available, another type of error is introduced in trying to fill in the gaps for those industries and years where estimates are not known. Methods such as interpolation, extrapolation, use of imputed weights, inserted trends, and "blowing up" of sample data are used in order to fill in missing data which introduce uncertainties of their own. Such gaps are particularly noteworthy for underdeveloped countries which produce only partial statistics or for countries as the Soviet Union where statistics of certain sectors are withheld for political and other reasons or could not be obtained because of the effects of war and revolution. This third source of error is, therefore, of great significance in international comparisons of national incomes and figures derived from them, such as growth rates, investment rates, etc.

4. Measurement of Error.

These three basic possibilities for error are present to a greater or lesser extent in each of the components. National income is a total of composites which differ in reliability from sector to sector and year to year, and hence the error of the composite is a "complex amalgam of errors in the parts whose magnitude is not easily determined."[1] The National Income Division of the Department of Commerce provides no measure of the possible error, taking the position that "meaningful mathematical measures of reliability cannot be calculated for national income statistics; only a frank evaluation of the sources and methods underlying them can provide the understanding which is needed for their effective use in economic analysis."[2] Any quantitative

[1] S. Kuznets, "Discussion of the New Department of Commerce Income Series, National Income: A New Version," op. cit.

[2] U. S. Income and Output, p. 48.

estimate is left to the user of the statistics based on his knowledge of the sources and methods as provided by the Income and Output and the 1954 National Income Supplements. The national income is built up of so many cells, and since there may be several types of error operating in each cell (for each cell may be the composite of many series), a variety of procedures must be used in compressing the economic activities of the nation into the accounting framework.

We have seen that conceptual differences play an important role in casting doubt upon the accuracy of any one statistic over the others that are available and over the previous unrevised estimates. However, as yet, no one has arrived at a measure of the margins of errors which are inherent in the estimates of national income. These margins could only be stated by the agencies that collect the basic data or the compilers of the aggregates. Since most of these groups either seem to have ignored the problem, or simply refuse to deal with it systematically, it becomes impossible for the user to determine with what confidence he may employ the data. The fact that little or nothing is said about accuracy[1] is more dangerous than if the margins of error were frankly stated to be very high. This is particularly important in view of the great and increasing importance attached to national income figures in policy making.

To throw the burden of estimating the errors and the reliability upon the user, though exceedingly convenient for the maker, is a totally inadmissible procedure. How can the individual user be expected to accomplish something where the government with its vast resources fails? This kind of evasion is also frequently encountered, as we have seen, in other

[1] British authorities, in charge of putting together national income statistics, have officially classified categories of quality (cf. below, S.). On the other hand, as I am reliably informed, the German Government has expressly forbidden that errors of components of national income statistics be indicated by the government agencies responsible for producing these data!

fields of government statistics. It either shows that one is lacking in clear ideas and procedures or does not dare to use them since they would show up the tremendous limitations of the figures which the government itself uses freely in the pursuit of its business. It certainly demonstrates that those who attempt to place the burden of proof on the reader or user have only an inadequate idea of proper scientific procedure.[1]

5. Direct Estimates of Error by Expert Judgment.

The most important study that has been made so far is by Simon Kuznets[2] on margins of error in national income estimates. It has pointed the way for future work. Kuznets considered the aggregate national income as composed of 520 cells (40 industries, 13 income and employment categories). Then he and two of his co-workers attempted to classify each of these entries according to its margin of error. The possible margins of error were grouped into four categories:

I.	5-10%	with average of	7.5%
II.	11-20%	" " "	15 %
III.	21-40%	" " "	30 %
IV.	41-80%	" " "	60 % ,

and for each cell each of these three investigators made independent classifications. An average was taken of their judgments[3] and the deviation between

[1] We might, at this point, recall the discussion of autocorrelation of errors (Part I). It is frequently maintained that national income statistics may be subject to a very considerable bias or that errors in these statistics are highly autocorrelated. Though this may well be true, there can be little doubt that national income statistics are subject to considerable margins of error which do not have this convenient property, and these errors must render small changes in the figures meaningless.

[2] S. Kuznets, National Income and its Composition, vol. II, chapter 12 (National Bureau of Economic Research, 1942).

[3] On the use of judgment in statistics and the construction of judgment indexes, cf. John W. Tukey, "Future of Data Analysis," Statistical

them was noted. As a result, a measure was obtained of the general magnitude of errors in each of the component estimates of national income as well as of the aggregate itself. (Estimates were judged both directly and by their component parts and the error for the direct estimate was in most cases noted to be lower because of cancellations.) From this classification Kuznets distinguished three groups of industries according to the relative margins of error judged to be present in their estimates: First, with a margin of error well below 15 percent (in categories I and II above), were the basic manufacturing industries and public utilities — electric light and power, steam railroads, street railways, telephone, telegraph; second, with margins of error of about 15 percent but well below 30 percent, were agriculture, mining, manufactured gas, pipe lines, trade, banking, insurance, and government — industries for which information is extensive but not complete; and third were industries with an error margin of about 30 percent and higher — construction, water transportation, real estate, direct service industries, and the miscellaneous division. Kuznets' estimates of these margins of error in the period 1919 - 1935 are shown in Table N-3, which also compares his results with the average importance of each industrial division in the period 1919 - 1938. (Note that even so meticulous an investigator as Kuznets computes a mean of pure guesses to two decimals.)

In the examination of margins of error in estimates of number of employed and engaged, the same grouping of industries was noted.[1] However, in this case the total margin, which was found to be 16 percent, was expected

---

Techniques Research Group, Section of Mathematical Statistics, Department of Mathematics, Princeton University (Princeton, N.J., July 1961). Cf. also D. R. Cox, "The Use of a Concomitant Variable in Selecting an Experimental Design," Biometrika, 44, 1957, pp. 150-158, for the use of judgment indexes as competitors for covariance.

[1] S. Kuznets, National Income and its Composition, pp. 191-97.

Table N-3

Net Income Originating, Margin of Error by Industrial Division, 1919-1935 and Relative Importance of Division, 1919-1938

Industrial Division	Margin of Error: Directly Estimated Mean	Directly Estimated Relative Deviation*	Estimated by Parts Mean	Estimated by Parts Relative Deviation	Percent of National Income	
Group I						
Electric light and power	11.43	32	12.36	34	1.4	
Manufacturing, total	9.45	36	9.76	17	21.0	
Steam railroads, pullman and express	7.50	0	7.50	0	5.4	
Street railways	10.98	29	11.06	30	0.74	
Telephone	7.50	0	7.50	0	0.94	
Telegraph	7.50	0	7.50	5	0.16	
Group total						29.64
Group II						
Agriculture	12.40	40	24.32	99	9.6	
Mining, total	13.10	45	17.94	29	2.2	
Manufactured gas	17.90	27	20.32	30	0.25	
Pipe lines	15.00	35	12.99	37	0.20	
Trade	20.50	62	24.82	31	13.5	
Banking	15.63	11	15.25	12	1.4	
Insurance	14.80	2	17.56	5	1.6	
Government, total	17.66	18	29.31	39	11.6	
Group total						40.35
Group III						
Construction	26.91	12	28.36	8	3.8	
Water transportation	27.27	14	26.26	17	0.73	
Real estate	36.78	35	38.33	25	8.9	
Service, total	27.27	14	39.10	9	12.6	
Miscellaneous	54.56	14	49.36	12	4.0	
Group total						30.03

* The relative deviation is $\frac{\sigma}{\bar{x}} \times 100$.

Source: S. Kuznets, National Income and its Composition, pp. 513-4; 166-7.

to be an exaggeration, since it was judged by parts and therefore did not permit cancellations to take place.

The weighted margin of error for the estimate of national income was found to be about 20 percent by this method of averaging "expert guesses" of the components, and summing. However, Kuznets felt that this figure was exaggerated, that if interest and dividends were included in nationwide estimates of income, and if entrepreneurial withdrawals were combined with entrepreneurial net savings, there would be substantial cancellation in the error margins assigned to each component separately. Also, if the statistics of employee compensation were examined for the nationwide total, their error factor would be considerably reduced. As a result, Kuznets infers that an average margin of error for national income estimates of about 10 percent would be reasonable.[1]

Although Kuznets' evaluations are little more than "informed opinions," which are based on studies of the errors involved in interpolation and extrapolation, comparisons with estimates from other sources, and revisions of official data from time to time, they provide a method of finding margins of error in this field where none had existed before. Since the data in national income studies are "partly a byproduct of administrative activity, partly a result of direct observation of complex phenomena without controls designed to reduce the variations observed, the best that we can do is to express an opinion in quantitative form."[2]

Kuznets' observations are of critical importance. The judgments and very carefully considered estimates of this eminent authority, in a field

---

[1] A very well known and highly informed American statistician has recently stated that a ± 20% error in U. S. national income statistics is not implausible. Unfortunately, the text of his speech has not yet been released for publication.

[2] S. Kuznets, National Income and its Composition, p. 55.

which he has helped so much to develop, touch on a class of economic statistics that are in wide use and employed for the most diverse purposes. The almost religious attention paid to "GNP" — it being continually used and quoted in the teaching of economics as well as in Government and in the business community — would lead one to expect that criticism would be reacted to sharply. This has not been the case. The textbooks on national income and macroeconomics show little if any evidence of awareness of these difficulties and limitations. The trade journals likewise go on accepting the statistics at face value and do not seem to be conscious of their severe limitations. This is a thoroughly unsatisfactory state.

The method described above may be very useful in future work in estimating margins of error. This is of particular interest for input-output tables which are, fundamentally, also accounts of national income. Input-output tables do, however, involve even much finer measurements. If they could be fully explained, they would give more information directly useful for economics than the also highly desirable national income figures.

It behooves us to pause in order to see what even a 5 percent difference in national income means. Taking the United States and assuming a gross national product of about 550 billion dollars, this error equals ± 30 billion dollars. This is more than twice the best annual sales of General Motors, the country's (and the world's) largest industrial corporation. It is far more than the total annual production of the entire electronics industry in the United States. Yet we have seen that a 10 percent error is even more reasonable: but that amounts to a plus or minus variation exceeding the entire defense budget of the nation, or it is about three times the total exports of the United States! The possible differences are, of course, not concentrated in the manner of these illustrations; instead they are scattered in an unknown way throughout all activities producing the

national income. On the other hand, the reader, like everyone else, has probably become conditioned to accept economic data as being so highly accurate that even a mere 1 - 2 percent variation of national income is considered significant enough for making statements about "true" variations in the state of the economy. Yet 1 - 2 percent in gross national product are 5 - 10 billion dollars, and even that is an amount which few would judge irrelevant for the economy of the United States. On the contrary these are amounts now used in order to estimate and predict the future performance of the entire economy and to justify far-reaching policy measures.

If it seems unreasonable to accept errors illustrated by such absolute magnitudes, the answer is that as always the burden of proof is on those who wish to continue using these (or any other) data in the traditional manner. Illustrations of the above kind are apparently needed, in order to stress the seriousness of the situation and to caution against the uncritical practices of the day, the mere mentioning of the percentages not having made the necessary impression. One should also recall the fact that the population census of 1950 for the United States failed to account[1] for the presence of approximately 5,000,000 persons. That is equivalent to the United States not possessing cities of the size of Chicago plus Detroit, which should certainly make some difference! And certainly the presence or absence of 5,000,000 people should make some difference in personal disposable income, or in gross national product, even if some of those omitted were children or infirm!

If, on the other hand, confidence in the published figures be maintained, i.e., the existence of errors or at least of errors of the above magnitudes be denied, then the procedures and methods of evaluating the statistics which have led to the above mentioned error estimates have to be

[1] Cf. A. Coale and M. Zelnik, A Study of White Births and Birth Rates in the United States, 1855-1954, and of Completeness of Enumeration in Decennial Censuses, 1880-1960. (Scheduled for publication in 1962.)

rejected. That, too, would be a very serious matter, since it would be tantamount to questioning large parts of present statistical theory. In the area of economic statistics some problems have arisen that are, as yet, not even dealt with adequately in current statistical methodology.

The decision whether to accept the official figures or the error estimates is ultimately a matter of intuition. If the error estimates were to run as high as 50 percent, we would probably reject the method of arriving at such estimates. But is this point already reached at 5 or even 10 percent? This is most doubtful, and we will therefore have to accept living with data which are widely thought to be much better. In particular we will have to accept the meaning of errors of this magnitude in the calculation of growth rates (cf. the following chapter).

To summarize: The rudimentary information obtained so far about errors in national income statistics shows that these important statistics are in especial need of decisive improvement. In particular, it should be stressed that the present exaggerated practical applications must be avoided. It is not unusual, for example, to consider changes in the national income figures of plus or minus one-tenth of one percent (or even less!) as significant for either theory or policy. In the face of the facts such procedure is completely void of meaning.

It is distressing to see that even the high-placed Council of Economic Advisors to the President engages in the practice of taking the figures for gross national product and national income at face value. In its entire history it appears never to have investigated their accuracy, and as a consequence draws wholly unwarranted "conclusions" from alleged one percent changes of these great aggregates. The same applies to their treatment of growth rates.

6. Income vs. Product.

Gross national product totals are derived in two separate ways. One method is to sum the income of all the factors of production, i.e., employee compensation, profits, rent, net interest, and add indirect taxes, capital consumption allowances, and several minor items. Gross national product may be viewed also as the sum total of all expenditures paid out, and in this case it is the sum of expenditures for consumption, private domestic investment, government expenditures, and net foreign investment. If the counting were accurate, the income and product sides would exactly agree, as the sum of all incomes paid out in a given period must be equal to the sum total of all income received in that period. However, since in fact the two sides never balance exactly, there is a reconciling item termed "statistical discrepancy" which appears as an item in the national income accounts (by convention it is entered on the debit side). If the procedure of obtaining estimates of income and product is free from bias, the statistical discrepancy obtained from estimating income and product over the years should behave like a random error of measurement. In a study made a few years ago, the statistical properties of this discrepancy were examined. "Since the GNP estimates are built up from generally independent sources, the discrepancy may be taken to represent the net result of the numerous forces which introduce errors in the estimates of the detailed components on each side of the accounts."[1] However, while the discrepancy indicates that the totals do contain errors, it is taken by the National Income Division of the Department of Commerce to indicate a lack of consistency between the two sides and not as an absolute measure of the errors therein. If the National Income Division finds that

[1] A. J. Gartaganis and A. S. Goldberger, "A Note on the Statistical Discrepancy in the National Accounts," Econometrica, Vol. 23, No. 2, April 1955, pp. 166-173.

the discrepancy is large or erratic in movement, the estimate is reexamined and attempts are made to trace and eliminate the source of the discrepancy as far as possible.[1] But even though improvements may be made, a residual discrepancy remains. Gartaganis and Goldberger did find some indication of patterns of temporal interdependence in the estimates in some cases, particularly in the pre-1954 estimates.

7. Absolute Size of the Estimates, Relative Changes and Revisions.

The problem of the accuracy of national income statistics may be viewed from two aspects. First, how good are the totals, i.e., what is the probability that the real national income figure falls within, say, plus or minus twenty billion dollars of the published estimate? Second, and more importantly from a practical point of view, is the question of the reliability of changes in direction (flow) of the various national income series from year to year or quarter to quarter.

(a) Absolute magnitude. The absolute size of national income depends primarily on the conceptual foundations of the measure (besides the effects of the problems due to sampling, interpolation, etc., alluded to above). We have mentioned above the complicated conceptual problems of imputation, measurement of the government's contribution to national income, inventory and depreciation treatment. We emphasize again that statistical measurement is only one part of the problem of the accuracy of national income statistics; conceptual difficulties prove to be no less plaguing.

The National Income Division laid the foundations for the present national income estimates in 1947, when revisions were made for the period

[1] For example, see the description of the revision of consumer expenditures in 1950 by the Department of Commerce, made when the discrepancy turned out to be a positive figure. The reexamination resulted in a scaling down of the initial calculation of consumer expenditures by 1½ percent of the 1950 total, or by $3½ billion. U. S. Income and Output, pp. 6-7.

1929-1946, incorporating the conceptual framework and statistical methodology which were established in 1947. In Table N-4 (A.) we see the magnitude of adjustments in the revision of national income estimates in 1947. On the average, the total revision for this period is ± 7.1 percent (this is for the entire period; we have chosen only a few years as illustrative) and the average statistical revision is ± 1 percent. Changes due to differences in concept account for the major part of the revision,[1] but it can be seen that in some cases the statistical revision was of some importance. In 1946 there was a 4.2 percent increase over the earlier estimate, due to statistical revision, and in 1932 there was a decrease of 2 percent. On the other hand, revision due to changes in concept produced changes of 13.1 percent (in 1944) and an average revision over the period of 6.2 percent. We can thus conclude that even though the revisions here are largely due to changes in concept, they do reflect the considerable measure of uncertainty which even their makers attach to these statistics. This is particularly noteworthy regarding the revisions in the 1940's, which are larger than earlier ones. Compared with much smaller earlier revisions, this shows that it is not generally true that more recent statistics are subject to less doubt than earlier ones (assuming that these revisions themselves can inspire confidence).

The years 1947-1960 show that the process of correction continues and that both positive and negative changes are required. In Table N-4 (B.) we show the differences between the preliminary and the latest estimates of national income in this period. The largest absolute change is for 1947, with + [illegible] billion (2.4%), incidentally a rather critical year for the currently used procedures in determining business cycle

[1] This statement is, of course, only plausible under the assumption that the basic, underlying figures which enter into different conceptual frameworks, are usable. Whether this is the case or not is a separate question.

Table N-4

Comparison of Estimates of National Income

A. Reconciliation of New and Old Series of National Income: 1947 Adjustments[1] 1929-1946 ($ billions)

1929-1934	1929	1930	1931	1932	1933	1934
National income (new)	87.4	75.0	58.9	41.7	39.6	48.6
National income (old)	83.3	68.9	54.5	40.0	42.3	49.5
Total revision	+4.1	+6.1	+4.4	+1.7	-2.7	-0.9
Amount of revision due to concept changes	+4.4	+6.7	+5.1	+2.5	-2.1	-0.6
Statistical revision	-0.3	-0.6	-0.7	-0.8	-0.6	-0.3
Total revision as % of old series	+4.9	+8.9	+8.1	+4.3	-6.4	-1.8
Revision due to concept changes as % of old series*	+5.3	+9.7	+9.4	+6.3	-5.0	-1.2
Statistical revision as % of old series*	-0.4	-0.9	-1.3	-2.0	-1.4	-0.6

1935-1940	1935	1936	1937	1938	1939	1940
National income (new)	56.8	64.7	73.6	67.4	72.5	81.3
National income (old)	55.7	64.9	71.5	64.2	70.8	77.6
Total revision	+1.1	-0.2	+2.1	+3.2	+1.7	+3.7
Amount of revision due to concept changes	+1.3	+0.2	+2.3	+3.1	+1.3	+3.6
Statistical revision	-0.2	-0.4	-0.2	+0.1	+0.4	+0.1
Total revision as % of old series	+2.0	-0.3	+2.9	+5.0	+2.4	+4.8
Revision due to concept changes as % of old series*	+2.3	+0.3	+3.2	+4.8	+1.8	+4.6
Statistical revision as % of old series*	-0.4	-0.6	-0.3	+0.2	+0.6	+0.1

1941-1946	1941	1942	1943	1944	1945	1946
National income (new)	103.8	136.5	168.3	182.3	182.8	178.2
National income (old)	96.9	122.2	149.4	160.7	161.0	165.0
Total revision	+6.9	+14.3	+18.9	+21.6	+21.8	+13.2
Amount of revision due to concept changes	+6.1	+13.2	+17.4	+21.1	+19.5	+6.3
Statistical revision	+0.8	+1.1	+1.5	+0.5	+2.3	+6.9
Total revision as % of old series	+7.1	+11.7	+12.7	+13.4	+13.5	+8.0
Revision due to concept changes as % of old series*	+6.3	+10.8	+11.6	+13.1	+12.1	+3.8
Statistical revision as % of old series*	+0.8	+0.9	+1.0	+0.3	+1.4	+4.2

Table N-4

Comparison of Estimates of National Income

B. Preliminary Estimates and Latest Revisions of National Income 1947-1960 ($ billions)

	1947	1948	1949	1950
National income (revised)[2]	198.2	223.5	217.7	241.9
National income (preliminary)[3]	202.6	224.4	221.5	235.6
Difference (bil. $)	- 4.4	- 0.9	- 3.8	+ 6.3
Difference as % of preliminary series	- 2.2%	- 0.4%	- 1.7%	+ 2.7%

	1951	1952	1953	1954
National income (revised)[2]	279.3	292.2	305.6	301.8
National income (preliminary)[3]	275.8	290.4	307.7	300.0
Difference (bil. $)	+ 3.5	+ 1.8	- 2.1	+ 1.8
Difference as % of preliminary series	+ 1.3%	+ 0.6%	- 0.7%	+ 0.6%

	1955	1956	1957	1958
National income (revised)[2]	330.2	350.8	366.9	367.4
National income (preliminary)[3]	322.3	342.4	358.0	360.5
Difference (bil.$)	+ 7.9	+ 8.4	+ 8.9	+ 6.9
Difference as % of preliminary series	+ 2.5%	+ 2.5%	+ 2.5%	+ 1.9%

	1959	1960
National income (revised)[2]	399.6	417.1
National income (prelim.)[3]	398.5	417.1
Difference (bil.$)	+ 1.1	0
Difference as % of preliminary series	+ 0.3%	0.0%

*These figures may not add up to the total revision because of rounding.

Sources:

[1]U. S. Department of Commerce, "National Income and Product Statistics of the United States," Supplement to Survey of Current Business, 27: 14, July, 1947.

[2]National income (revised) from Survey of Current Business, July, 1961.

[3]National income (preliminary) from Survey of Current Business, Annual Review numbers, every February issue from 1948 to 1961.

turning points. If the change finally approaches zero (for 1960) this should not be taken as a contradiction to the last sentence in the preceding paragraph. The reason for the small percentage changes of the last 2-3 years is simply that it sometimes takes ten years before the final figure is arrived at, as will be seen from Table N-2 where the consecutive corrections are listed.

The problems of revision which are so important in these measures, although of course affecting the absolute size of the magnitude of the various income and product series, will be discussed in the next section. There we shall deal, first, with United States data and practices, and second, with the experience in the United Kingdom, in order to show that the problem presents itself everywhere when national income statistics and similar accounts are being put together.

Revisions of previously reported figures may derive from conceptual changes (but there have been no significant alterations in concept since the major revision of 1947), from later or more reliable bench marks, and from better statistics or advanced methods of processing.

(b) Relative changes and revisions. The fact that the estimates of the absolute level of national income may leave us with doubts as to their reliability does not necessarily mean that the flow figures or changes from period to period are subject to the same magnitude of error. For example, the conceptual difficulties of imputation would present no problem when calculating short-run changes if they are consistently handled; it is only as a consequence of a change in a procedure or concept that the flow figures would suffer. However, this does not mean that the user of the statistics on change has reliable estimates of movements in the national

---

[1] A Critique of the United States Income and Product Accounts, op. cit., p. 291.

income series. Even if we consider only the revisions which are regularly being made, the figures are always of a tentative nature.

We will discuss the problem of revision and changes in the quarterly estimates published by the National Income Division. Few series are more widely used in analyzing the nation's economy than the quarterly estimates of national income and gross national product (and the monthly series on personal income).[1]

There are three basic criteria by which to judge the reliability of quarter-to-quarter changes of the national income series. The first is the extent and nature of "bias," i.e., the extent to which the initial estimates tend to be too high or too low on the average. It is necessary to point out that by measuring the amount of bias using the final estimate as reference, there is the implicit assumption that the final revised estimates are "correct." Only for purposes of comparison can the final estimates be considered correct in the sense of being at least the most correct figures available. In other words, it is usually considered that progressive revisions get us closer and closer to the actual truth. The second is a measure of the extent of the average revision, i.e., a measure of the firmness of a given quarter-to-quarter percentage movement. Third, one may consider the proportion of the times the first estimates of change fail to give the "correct" direction of movement.[2] But what is the "correct"

---

[1]Throughout this section we do not mean to imply that users of any of the series under discussion, gross national product, national income, etc., would confine themselves to any one estimate as an all-purpose indicator of the state of the economy. Judgment as to direction or magnitude of change in the economy, or in any part of it, should rest on a number of indicators, e.g., the Federal Reserve Board index of production, employment rate, steel production, etc. Of course, each of these pieces of information is subject to its own unknown errors and there is no guarantee that they would, in general, cancel each other out.

[2]Arnold Zellner follows much the same lines in a study of the provisional and revised quarterly estimates of gross national product and its components in the period 1947 to 1955. One of his most striking findings

direction when successive revisions of the figures for the same year sometimes are positive, sometimes negative in succession, as has repeatedly been the case? The latest figure for a given year is not always the final figure even for several years after the given date. Hence there is not necessarily a value towards which the successive, repeated corrections clearly converge. Only if this were known to be the case could one speak of "correct" directions, and if one could be sure of this, the limiting value could easily be computed.

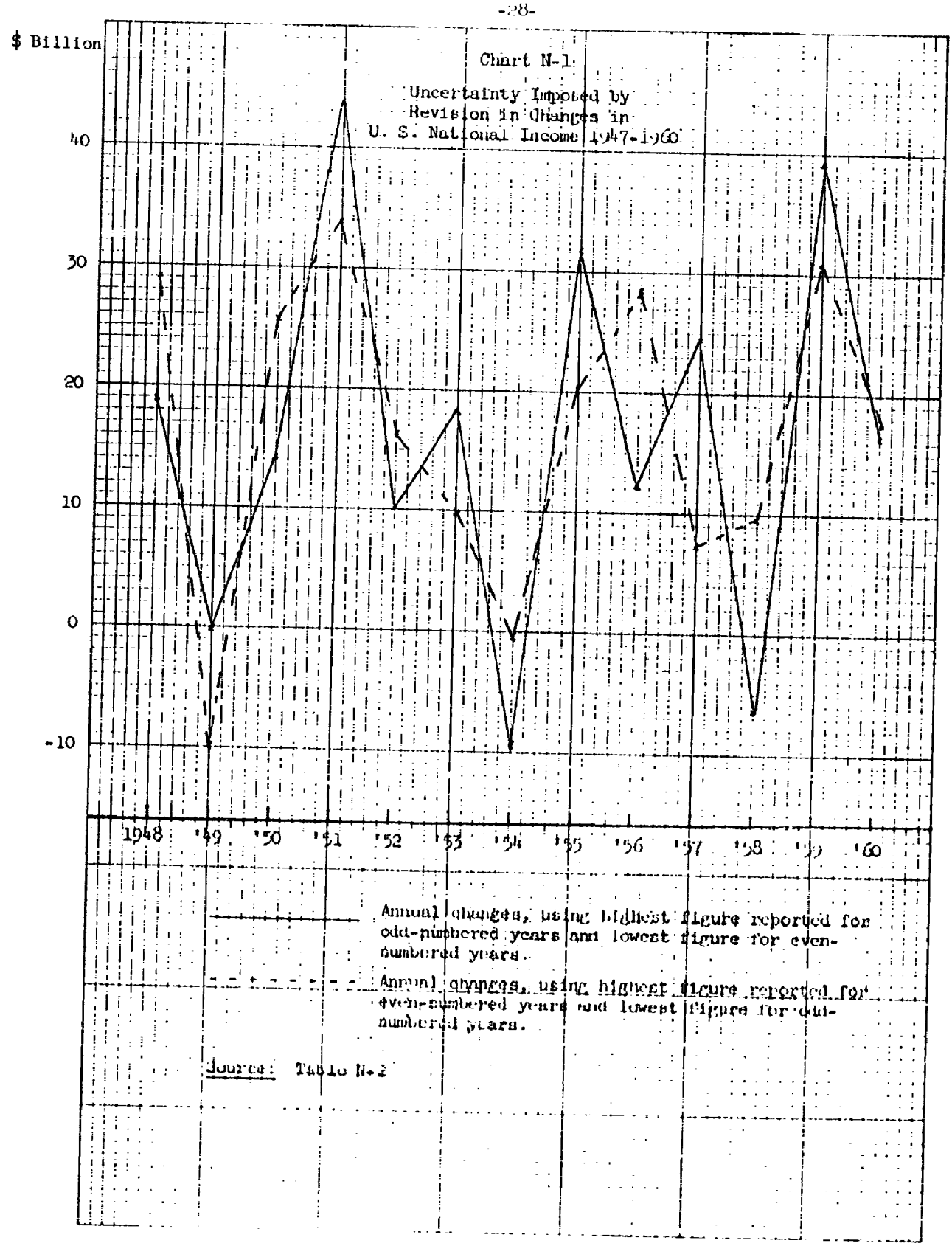
Table N-2 shows the repeated revisions of United States National Income since 1947, and Chart N-1 represents some implications. We see, as already mentioned, that revisions of the originally given figure sometimes still come after ten years! We also note that there seems to be no tendency for the number of revisions to decrease significantly. It is, of course, desirable that revisions be made when new information becomes available. But their frequency and long delays betray an uncertainty — no doubt justified — permeating the whole field, which is in striking contrast with the widespread immediate use of the first given figure and the assumption that it is significant to one billion dollars or less. We dispense with a more detailed description of this table since it offers no difficulties. But we emphasize again that each consecutive figure, i.e., revision for the same year, is afflicted with its own error and that the mere fact of a revision as such offers no guarantee whatsoever that the error has thereby been reduced.

---

was: " ... whereas provisional estimates of GNP disagreed with revised estimates only in 9 cases, these occurred at the lower turning points of the 1948-9 and 1953-4 recessions." (And this was all before deflation by doubtful price indexes!) Cf. Arnold Zellner, "A Statistical Analysis of Provisional Estimates of Gross National Product and its Components, of Selected National Income Components, and of Personal Saving," Journal of the American Statistical Association, 53, March, 1958, pp. [illegible].

Chart N-1

Uncertainty Imposed by Revision in Changes in U. S. National Income 1947-1960



Source: Table N-2

Neither is the converse the case: a lack of revision does not necessarily imply that the given figure is good (that is particularly to be borne in mind when judging components of aggregates!). More often than not it simply means that no revision has taken place. Since no error is stated by the makers of the statistics in the first place, nothing is being said about any likely changes in the error either. The reader, of course, has no way of determining these changes for himself.

Chart N-1 shows what paths of description of change in national income could have been chosen, assuming that each of the different values given for a certain year is equally likely and would actually have been reported instead of the whole set of figures given for each year. This chart expresses clearly some of the uncertainties prevailing in this field, though — again — nothing is shown about the far bigger uncertainty residing in the basic error of each number. If that were expressed it would demonstrate that the path, leading from year to year, can vary enormously, and that only the very broad tendency of an increase of national income over longer periods of time can be asserted with confidence. These observations, incidentally, should be viewed as casting serious doubts on the usefulness of national income figures for business cycle analysis. The idea that quarterly, let alone monthly, figures of gross national product, national income, etc., could be obtained, even with the most modern recording devices, without appreciable error, is nothing short of grotesque. And when we do have error and try to determine growth rates, we are by necessity exposed to the consequences described by the computations in the subsequent chapter on Growth Rates.

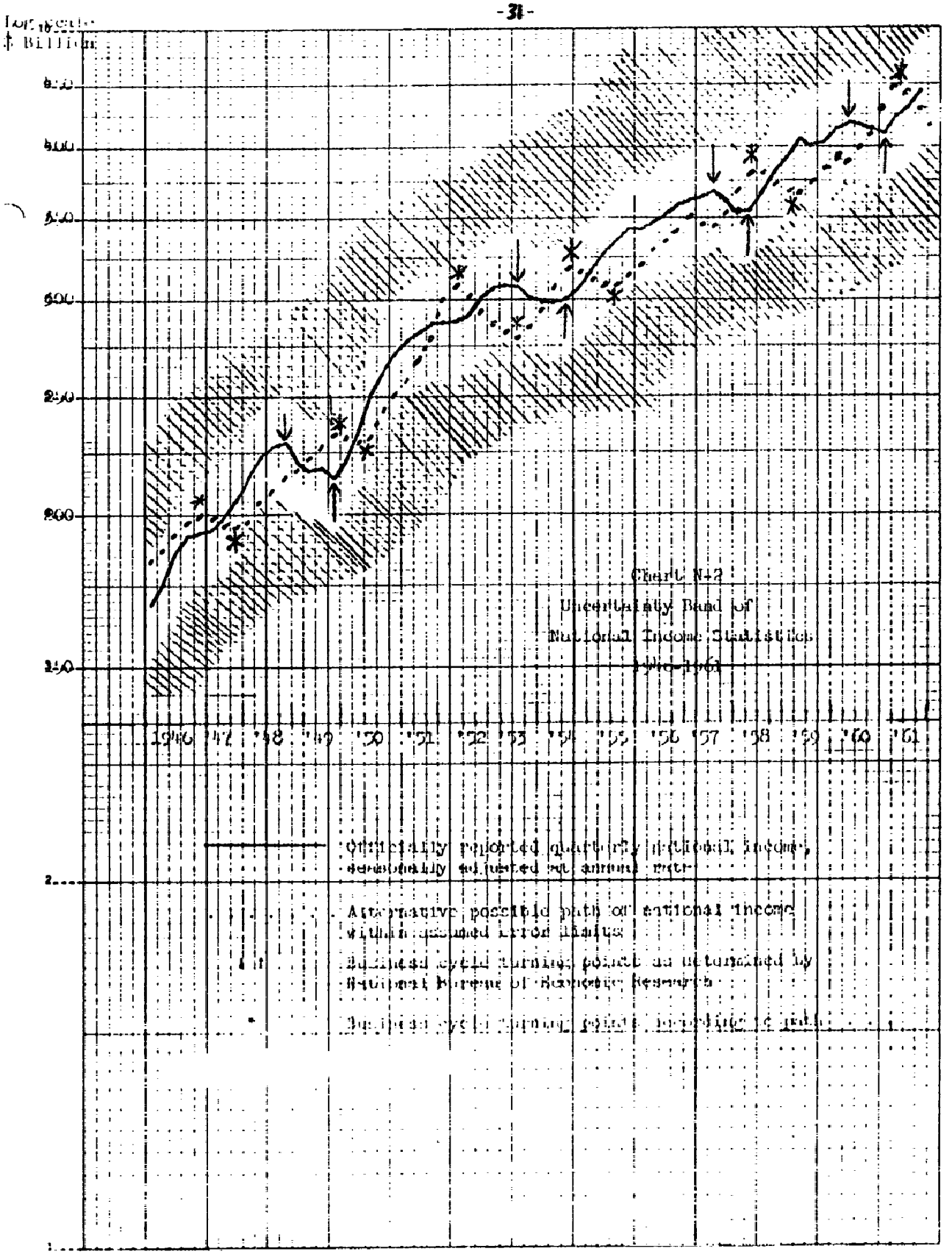
Similar statements apply to British data (cf. section 4 below). If these experiences are at all [illegible] for the making of national income statistics, then one can form an idea what ups and downs the

revisions of other, perhaps less advanced, countries show or should show, and how uncertain it is to attribute to any one year a "true" figure, where "true" means merely conforming to the existing method with its known limitations and its unknown error.

Chart N-2 shows the implications of ± 10 percent errors in the seasonally-adjusted quarterly national income figures. Between the shaded areas lies the zone of uncertainty surrounding the data. No single quarter-to-quarter change is significant within this band. Changes persisting for several quarters are seldom significant. By and large the national income figures contribute little confirmation of the post-war business cycle, its turning points shown by the small arrows, as the different possible profile of national income shown by the dotted line brings out, its turning points marked by stars. This alternative path is not in any sense "correct"; but the point is that we do not know where in the region the statistics should lie.

In Table N-5 the various revisions are shown as related to the principal component series; there is, clearly, an interest in observing which categories of economic activities contributed most or least to the revisions.

Preliminary estimates for most series of the quarterly national income series are made available in the Survey of Current Business in the issue of two months after the end of the quarter; for corporate profits and national income, the lag is an additional two months. Subsequently, however, there are still sizeable revisions in the published series. The quarterly statistics are based generally on more limited information than the annual series. First revisions are made three months after first publication, and in the July issue of the Survey of Current Business the data may be revised for the last two or three years. The three-year period allowed for possible revision permits the incorporation of data accumulated by such annual undertakings as the Statistics of Income of the Internal



Officially reported quarterly national income, seasonally adjusted at annual rate

Alternative possible path of national income within assumed error limits

Business cycle turning points as determined by National Bureau of Economic Research

Chart N-2

Uncertainty Band of
National Income Statistics
1946-1961

Table N-5

Selected National Income Series, 1947-1958:
Summary Measures of Quarter-to-Quarter Percent Movement

National Income Series	Average revision as a % of average movement of revised estimate of change	% of time direction of movement missed	Average bias in first estimates of quarter-to-quarter movement
Compensation of employees	26 %	6 %	- .17 %
National income	26	4	- .18
Personal income	28	6	- .15
Personal consumption expenditures	34	2	- .31
Gross national product	38	11	- .22
Corporate profits before tax	40	19	.53
Gross private domestic investment	61	19	.01
Proprietors' and rental income	96	28	.05

Source: "Revisions of First Estimates of Quarter-to-Quarter Movement in Selected National Income Series, 1947-1958 (Seasonally Adjusted Data," Statistical Evaluation Reports, Report No. 2, Office of Statistical Standards, Bureau of the Budget, February 1960, p. 23.

Revenue Service, various biannual censuses, etc.

The magnitude of the revisions may give some measure of the "firmness" of the particular component series of national income. A series with smaller quarter-to-quarter revisions does not necessarily indicate greater accuracy. It may mean simply that no better data have become available, and that the original data, however weak they may be, remain the best available. On the other hand, a series with large revisions certainly indicates that the original estimates were weak. As a practical matter, series which are usually considered among the better

series have the least relative revision.[1] In this sense, national income, compensation of employees, and personal income are "good" series, whereas proprietors' and rental income and even corporate profits are relatively weak series. Observe, however, that corporate profits are based on (audited!) balance sheets (with all their limitations as discussed in Part I), and income tax returns, while for personal income not even the number of persons receiving income is reliably known (see reference in footnote 1, page 18).

In Table N-5 the selected national income series have been grouped, using as criteria the relative revision of the first available estimates of change, the percent of times the first estimates missed the direction of movement, and, to a lesser extent, the average bias, which is generally insignificant in these series.

For most series the bias is small and therefore not significant, with the possible exception of corporate profits where initial estimates have averaged high. This series, so widely used, especially by security analysts, trying to determine the prospects of individual stocks, is one of the most difficult to measure with assurance. One of the principal difficulties is associated with making proper allowances for depreciation, which affect the magnitude of profits and are taken directly from income tax returns, no matter what method is used. In addition, this series has been revised on the basis of new data and by improvements in the processing of data.

The average revision is a measure of the dispersion of the original estimates of quarter-to-quarter percent change about the corresponding revised estimates. For the relatively "firm" series such as compensation

---

[1] The discussion in these paragraphs is based for the most part on "Revisions of First Estimates ... ," Ibid. (Note that a "seasonal adjustment" of data of this kind is a most questionable operation, mechanically applied but devoid of meaning.)

of employees and national income, the average swing from quarter to quarter is about two percent and the average revision is about .5 percent, or about one-fourth of the quarter-to-quarter movement. For "weaker" series such as rental and proprietors' income, the average revision actually is about as large as the swing itself. In this case, then, the measure of change is not particularly meaningful. Since there is no bias, the direction of the revision is completely random; therefore, it is just as likely as not that the observed change will be completely washed out when new revised data are published. The point is, of course, very significant for the user of these data: "In current business analysis, it is of great importance that the "truth" be as closely approximated within the current period as possible. Revision a year later, while significant for historical purposes, comes too late for the analyst in his diagnosis and prognosis of current trends."[1]

Thus, in series where the average revision is large relative to the quarter-to-quarter movements, the first estimates may often fail to give the "correct," i.e., ultimately asserted, direction of movement. For example, in the period 1947-58, proprietors' and rental income failed to detect the direction of movement 28 percent of the time and corporate profits 19 percent of the time. This may be compared with national income, which missed only 4 percent of the time, and compensation of employees, which was off course 6 percent of the time.

In summary, we conclude that while most United States national income series are relatively free from bias, there are large differences in the firmness of these series. When reliable estimates of the direction in which the economy is moving are needed and when such estimates are to be obtained from national income series, the firmer series should prove to be

[1] M. Cohen and M. A. Cairnburgh, "The Income Side: A Business User's Viewpoint," in A Critique of the United States Income and Product Accounts, op. cit., pp. 191-2.

the better indicators. On the other hand the firm series are not always very important for more specialized purposes, since some of them, such as national income, are highly inclusive and so global in character that they can only figure in very aggregative economic models of low power of resolution. Corporate profits and gross private domestic investment on the other hand would be very interesting for estimating, say, future activity on the stock market, but they are definitely weak series and therefore of little use when needed.

8. British National Income Statistics and Revisions.

The high level of British statistics and the pioneering work in the field of national income associated with such names as A. L. Bowley, Lord Stamp, and R. Stone, to mention only a few, does not remove the fact that British data are also of uneven quality and subject to important revisions with the associated uncertainties. The Central Statistical Office has frequently warned of the inherent unreliability of their estimates of national income.

In 1956 a system of reliability gradings was worked out[1] for the different components that appear in the various tables and accounts. There are three gradings: A grading of A indicates (with 90% confidence) that the reported figures are correct subject to a margin of error of less than 3% in either direction; B indicates an error of 3% to 10% in either direction, and C indicates that the error is greater than 10%. (This classification should be compared to the direct estimates by Kuznets, page 15, above.

The items receiving the better ratings correspond in some ways to those which in the United States were subject to the least violent revisions

[1] National Income Statistics: Sources and Methods (H. M. Stationery Office, 1956).

(cf. Table N-5, page 32). For example, consumers' expenditures, gross national product, wages and salaries received, received a grade of A, while profits, rent, income for self-employed received a grade of B. Note, however, that, as stated before, a lack of revision is precisely that and does not necessarily reflect reliability or trustworthiness of the estimate!

The nature and consequences of the British classification and the implications of the revisions of British national income figures were analyzed in an excellent study by Harry Burton.[1] This author has clearly and forcefully shown the magnitude of the errors of measurement when yearly revisions are considered, even for items that have an A rating. It is noteworthy that neither Burton's paper nor the original Blue Book has found the kind of reaction among economists and the public that they deserve. Instead, in Great Birtain as in the United States and elsewhere, national income statistics are still being taken at their face value and interpreted as if their accuracy compared favorably with that of the measurement of the speed of light.

Table N-6 summarizes the size of changes in estimates (in million £) for selected items for 1954 and 1955 when their corresponding errors are considered. There are two grade A items and one each for grades B and C. The official change is given in column 3. The changes in column 4 assume that the estimate for 1954 was too high by 1½% for grade A items, 6½% for grade B and 15% for grade C (there was a negative error for 1954) and that the estimate for 1955 was too low by corresponding amounts (a positive error for 1955). Column 6 gives the changes assuming the same magnitude of errors for grade A, B and C items, but this time assuming they are opposite in sign, i.e., that the 1954 figure is an underestimate and the 1955 figure

[1] Harry Burton, "The Reliability of National Income Statistics," Accounting Research, July 1957, pp. 246-251.

Table N-6

Changes in Selected Items: 1956 Blue Book (U. K.) 1954-1955
Low and High Gradings Error*
(£ million)

	Grade	Blue Book estimates		Change 1954 to 1955	Change: 1954-1955 with 1954 negative 1955 positive		1954 positive 1955 negative	
		1954	1955		Low error	Higher error	Low error	Higher error
	col.	1	2	3	4	5	6	7
Consumers' expenditures	A	11995	12783	788	1159	1407	417	169
GNP (at market price)	A	17964	19058	1094	1649	2019	539	169
Gross domestic capital formation	B	2624	3165	541	918	1119	164	- 37
Investment income due abroad	C	604	591	- 13	167	286	- 193	- 312

*Low gradings error: A ± 1½% B ± 6½% C ± 15%

High gradings error: A ± 2½% B ± 10% C ± 25%

Low negative bias for item in group A therefore means error of - 1½% .

Source: Burton, ibid.

---

an overestimate. Columns 5 and 7 correspond to 4 and 6 respectively, except that here the errors are assumed to be larger (2½% for grade A, 10% for B, and 25% for C).

Brief study of the table will indicate the tremendous uncertainties involved. Even for grade A items with low gradings error, the estimate of year-to-year change is very poor indeed — for consumers' expenditures a reported change of 788 might just as easily be as low as 417 or as high as 1159. For a high gradings error the range is 169 to 1407! Clearly, we can

say very little indeed for forecasting purposes.

For grade B and C items the situation becomes worse. Gross domestic capital formation (grade B), an extremely important item for business cycle analysts, was reported as increasing by £ 541 million from 1954 to 1955. Assuming a high gradings error, this figure could have been as low as £ -37 million[1] and as high as £ 1119 million. In other words, from the figures alone the economy could have been in a major depression or in a superboom!

British national income statistics are subject to frequent and wide revisions, as are those of the United States (cf. Table N-2). Table N-7 gives the figures for selected items for 1952 as reported in 1953, 1954, 1955, and 1956. While the revisions for grade A are minor and generally go in the same direction, the opposite is the case for grades B and C, where increases of the original estimate are followed alternatingly by increases and decreases. The net income from abroad — a very important item of the British balance of payments — underwent the most drastic revisions, a situation which leaves the user of these statistics with a great deal of uncertainty.

9. International Comparisons of National Incomes.

If the great difficulties in making reliable national income statistics for the United States and the United Kingdom are a good indication of the problems any country runs into, then we can infer a great deal about the value, or rather the lack of value, of international comparisons

[1] A negative gross investment figure is of course conceptually impossible and is only the result of arithmetic illustration. Zero gross investment would imply that there were no expenditures on capital formation and that the capital stock was decreasing at a rate equal to depreciation estimates. Clearly this is the extreme and is not even characteristic of a major depression.

Table N-7

Revisions of British National Income Statistics
for the Year 1952
(£ millions)

	Grade	As reported in: 1953	1954	1955	1956
Consumer expenditures	A	10478	10440	10570	10582
GNP (factor cost)	A	13653	13738	13861	13928
Gross capital formation	B	1931	2116	2089	2158
Net income from abroad	C	128	139	114	93

		Percent and direction of revision from original (1953) report		
		1954	1955	1956
Consumer expenditures	A	- 0.37	+ 0.87	+ 0.99
GNP (factor cost)	A	+ 0.62	+ 1.52	+ 2.01
Gross capital formation	B	+ 9.58	+ 8.18	+ 11.75
Net income from abroad	C	+ 8.59	- 10.94	- 27.35

Source: Burton, op. cit.

in this area. Such comparisons are freely made and far-reaching consequences are drawn, for example, when the different degrees of welfare, economic development, etc., are being evaluated. Many uses are strictly of a political character. Although attention has been paid,[1] especially to the conceptual

[1] Cf. especially I. B. Kravis, "The Scope of Economic Activity in

difficulties in making comparisons, there apparently exists no thorough exploration of the enormous obstacles in the way to meaningful results. We can only add a few remarks at this occasion.

There are two principal questions involved: first, one referring to the comparability and applicability of concepts (essentially developed in advanced industrial countries), and second, one referring to the quality of the component data — our principal concern — no matter what might be the answer to the first question. The main fact is, as far as concepts are concerned, that the organization and development of nations differ so widely that for each class or category of countries, to some extent, conceptually quite different situations arise. Concepts have been well analyzed in western industrial nations, but they have not been settled as the many revisions of data due to conceptual changes prove. They are less understood for more agricultural countries and for those which have incomplete monetary organizations, lack large markets, and have no all-pervading price system, and thus perhaps are oriented less towards pecuniary pursuits. There are many points in common, some of which were discussed above, such as the difficulty of accounting for work done at home without pay and for the same work performed against money payment. What is negligible in one kind of country can be important in another. This applies in particular to home-consumed agricultural produce, which is an enormous part of the total in underdeveloped agricultural countries and practically irrelevant in the United States. Clearly, this is far more difficult to measure in the former than in the latter; yet U. S. agricultural statistics are far from satisfactory. How, then, could the agricultural income of, say, Ceylon, the Congo, China, Bolivia, or Tibet be known at least as accurately? How can they be made

---

International Income Comparisons, in Problems in the International Comparison of Economic Accounts, Studies in Income and Wealth, Vol. XX (1957), p. 349 ff., as well as the discussion by E. E. Hagen and J. Viner. Cf. further P. Studenski, op. cit.

comparable, e.g., on a per capita basis, when even the number of inhabitants in these countries is in far greater doubt than in the United States?

Similar to hidden unemployment — that plague of meaningful unemployment statistics — there is hidden income which is probably the greater (on a percentage basis) the less developed the country is and the warmer is its climate.[1] Therefore, the numerical (money) expressions of income are in grave doubt as far as comparability is concerned. This is quite apart from the difficulty of converting different monetary units into a standard reference unit, especially when at the same moment of time some countries are undergoing inflation, others deflation, some have multiple exchange rates, others have virtually no foreign transactions, etc. The trouble of having to correct for changing money values exists, as we know, even for comparisons over time within the same country. Then it becomes important to realize that the deflation of national income figures is dependent upon the fact of whether it is done for aggregates or for components separately, when all have different degrees of reliability.

International comparisons are constantly being made. No doubt some information can be had from existing figures, and whether they are useful depends, as we shall not tire to repeat, on the purposes of the comparisons. To ascertain in a rather general manner the fact of gross differences of the income of different nations, to show that they differ by large factors,[2]

[1] Sometimes in a "negative" sense, i.e., in warm climates, certain efforts are not required and therefore no corresponding income is generated. For example, even with a U. S. per capita income, one would not heat a house in the tropics (though one may want to cool it) or wear fur coats. Therefore, the financial inability to provide heat and furs is irrelevant in determining the meaning of tropical income levels. This is partly a conceptual matter; clearly it is a different matter regarding the ability to provide food or education.

[2] But probably not by as large factors as is suggested by the official statistics. As Kuznets has observed several years ago, if the low figures were correct the inhabitants of the poorest countries would all have starved a long time ago.

and to see whether these differences have changed over the years, etc., is one thing, but to believe that we can state this and much more reliably to two, three, or even four "significant" digits is an entirely different matter. On the basis of the discussion of the figures of the United States and the United Kingdom, we could not accomplish the latter even for these two advanced and "similar" countries. Yet we need only to look at numerous United Nations publications to see that this is being done for the whole world without any further excuse. The most startling use — or rather abuse — is for determining allegedly comparable growth rates for different countries, on the basis of which far-reaching policy decisions are made. (Cf., the chapter on Growth.)

The Office of Statistical Standards of the U. S. Bureau of the Budget has published a "Memorandum on International Statistics,"[1] rating countries in four groups (for 1956): I = Very Good, II = Good, III = Fair, IV = Weak, as follows:

Table N-8

Accuracy of Statistics:
Various Regions

	All	I	II	III	IV
Number	64	17	9	18	20
Continent					
Africa	9		1	2	6
America, North	11	2	1	5	3
America, South	8		2	4	2
Asia	15		2	5	8
Europe and Oceania	21	15	3	2	1

Source: loc. cit.

[1] U. S. Congress, Joint Economic Committee, 85th Congress, 2nd. Session, 1958.

Only 17, about one-fourth, qualify as "very good." The reader will remember that the best is probably the United States and that for this country the average error is probably in the order of ten percent. One wonders what it might in fact be for the 20 countries where the statistics are "weak"? (Note: Not called "poor," or "bad"! Note also that an identification of the various countries was carefully avoided. This delicacy was perhaps dictated by the desire to avoid international incidents?)

The difficulties in this area of comparison are truly stupendous, and the many authors who have contributed so much to analyze and overcome them deserve praise and encouragement. There is no doubt that a gradual clarification of issues is taking place. Ingenious systems of national accounting have been worked out and are being widely used. There is a learning process under way that can bring beneficial results to those who participate in it.[1] But this is a slow process. And no matter how much the schemes and models improve, at the root of all trouble is the question of the basic statistics which are being put into a gigantic agglomeration. In a sense we see here something similar to the process of increasing the gap between the "have" and "have not" nations: The statistics of those countries which produce the better figures continue still to improve, while those of the poor countries improve at a much slower rate, in spite of assistance given to them by the United Nations and other agencies. These can only help with methods and concepts, but not with the collection of the basic data, where the root of the trouble lies.

A special problem is offered by the Soviet Union. The statistics of that country are exceedingly difficult to assess, but it is generally known that they are seldom what they purport to be. This is in part due to

---

[1] Vol. XX of Studies in Income and Wealth, quoted above, bears eloquent testimony to this fact.

the highly centralized administration of the country, a very different conceptual structure in the thinking about economics, the absence of a true price system, the immensity of the country (that is split into many semi-autonomous republics), etc. There has been a great deal of deliberate doctoring of statistics at many levels, for example, in order to make production results appear better than they were or to receive assignments of raw materials that would not otherwise be allocated, etc. Even Khrushchev has repeatedly referred to falsified accounts of various activities, especially in farming, and there is no reason not to assume that this was different in the time of Stalin. A particular trouble in measuring aggregates is — as in all other countries — the double counting, or rather the multiple counting. This becomes the more serious the more complicated the final products are, which is undoubtedly the case everywhere under the impact of the present industrial-scientific revolution. Double counting has apparently been a most serious defect of Soviet statistics, with the necessary result that accounts of national income have been exaggerated and increasingly so in more recent times (for the reasons just given). This is the upshot of criticism by S. G. Strumilin, a well-known Soviet economist.[1] For example, in 1945, industrial output was, according to him, more than 30 percent below 1940, rather than only 8 percent, as the official statistics show. Similarly, again according to Strumilin, industrial production rose from 1945 to 1956 only three-fold , rather than four-fold, as officially asserted. (This may, of course, still be an exaggeration, and is subject to our observations on growth rates, cf. below.) Though industrial production is not identical with national income, it is a substantial component; its difficulties are illustrative for the larger aggregate and show how limited the value

[1] Ocherki Sotsialisticheskoi Ekonomiki SSSR (Essays on the Socialist Economy of the USSR), Moscow, 1960.

is of any "growth factor" based on such data.

We do not propose to discuss the Soviet national income statistics here any further. They are analyzed by Studenski, where the international literature on this subject is mentioned.[1] It is clear that the arbitrariness in valuation, partly due to the attempted use of the obsolete Marxist labor value theory, is greater than in any country possessing a functioning price system. The difficulties exceed those in underdeveloped countries which also have a small market sector only; but there the market performs its true function of determining the allocation of resources, though it may not reach far into the economy. There can be no statistic of national income without measuring or postulating money streams, since there must be a common denominator and no one has as yet found a substitute for the money measure. That is why the Soviet national income is also given in terms of rubles, though what the latter mean is obscure: there is no free market and no free exchange rate at which to make comparisons. Apart from the assignment of monetary expressions on the basis of assumed, imputed values, there is still the underlying problem of the accuracy with which the physical phenomena can be recorded, whose values are to be expressed in monetary units. To count cows, to weigh the harvest, to enumerate machines, etc., is the same problem in all countries. Some solve it better than others. Most works dealing with Soviet national income pay little attention to these data problems, but more to the valuation question. Yet in Russia the former encounter even greater difficulties than in the United States or the United Kingdom.

[1] The most important and authoritative work on Soviet national income is the recent book by A. Bergson, The Real Income of Soviet Russia Since 1928 (Cambridge, Mass., 1961). This work takes up especially the above-mentioned difficulties regarding valuation, as well as many others, and attacks them with great ingenuity and competence. But the basic difficulties of the data remain, and our results concerning the lack of validity of commonly computed growth rates and their international comparability apply fully (cf. the following chapter).

Summarizing, we can state that statistics giving international comparisons of national incomes are among the most uncertain and unreliable statistics with which the public is being confronted. The area is full of complicated and unsolved problems, and in spite of the great efforts to overcome them, the progress is slow. This is a field where politics reigns supreme and where lack of critical appraisal is particularly damaging.

THE VALIDITY OF

ECONOMIC GROWTH STATISTICS AND GROWTH RATES

Contents

1. Introduction.

In recent years there has been much concern among economists, politicians, and the general public about the rate of economic growth of the United States and other countries. In addition to the goals of maintaining a high level of employment and providing for general stability in the price level, a third goal, that of maintaining a satisfactory rate of economic growth, has been added to the responsibilities of fiscal and monetary authorities.

The increased emphasis on economic growth has two reasons: first, the desire to provide a higher standard of living for our citizens; second, to determine how effectively the Western world competes in the cold war with the Soviet Union.

In general we have made it the rule in this book to concern ourselves chiefly with primary data and to stay away as far as possible from indexes, uses of data, and further conceptual issues. It became clear that this is not always possible, because the greater the inevitable aggregates are, the more important become concepts. Similarly, no index construction is trivial, and especially those involving the notion of "costs of living" and their comparisons over time and among nations require a descent into considerable depths of economic theory and mathematics. And when classifications are the issue, concepts again assume great significance. In discussing national income even such elusive notions as "community welfare" had to be touched upon. All these reappear by necessity when growth rates are mentioned, and therefore have to be given proper, though brief, consideration in what follows.

In these pages we will therefore comment on the meaning of economic growth, the problem of the accuracy of growth statistics, and the difficulties confronting the user of such information. Finally we shall discuss the

problems of comparison of growth rates for different periods and for different countries during the same period.

2. The Concept of Economic Growth.

Economic growth is generally considered to be the increase in the "real" output of the economy over time. There is controversy about the best index of economic growth, which is not surprising since "growth" is a very complex phenomenon, wherever it is encountered. Many models of the economy have been suggested, showing growth of different kinds, and business cycle theory has been concerned with the "expansion" phase of a business cycle as distinguished from the "expansion" due to long term growth. These ideas will not be discussed here. But reference should be made to the work of biologists and others who have studied growth of organisms[1] where it is shown that growing and dying processes are closely related within the same individual, a fact that corresponds somewhat to the frequent changes in technology used in economic activities. This often takes the form that one activity declines (e.g., railroads) while others expand (e.g., motor cars). The question immediately arises whether these two tendencies compensate or whether one outweighs the other, so that while there are some declines and some expansions, one could nevertheless discern a clear overall tendency. Since conflicting tendencies develop in many areas simultaneously, the difficulties of assigning weights are compounded. It is clear, at any

[1] Sir d'Arcy Wentworth Thompson: On Form and Growth (Cambridge: Cambridge University Press, 1948). Galileo has already shown in his Discourses that certain types of mechanisms, among them the bone structure of the human body, cannot be linearly extended at will. The idea of "broken economic trends" and the need for developing a theory of trend change (i.e., the interrelation of simultaneously rising and falling trends) is advanced in O. Morgenstern: Wirtschaftsprognose, Eine Untersuchung Ihrer Voraussetzungen und Möglichkeiten (Vienna: Julius Springer, 1929).

rate, that a complicated process of valuation is involved even when only the simple, current method for determining growth rates is used. This method is to take the movement of gross national product, adjusted for changes in the price level, as a satisfactory indicator. This rather crude procedure could be refined, for example, by taking national income or personal income. The more we refine the measure, the greater are the conceptual problems and the more difficult the decisions. In view of the important purposes of the determination of growth, the best possible measurements are barely good enough and any evasion of the underlying issues is inadmissible. Yet here we are only concerned with one single aspect, which is the reliability of the figures expressing the currently used measures, whatever their conceptual justification.

Gross national product, national income, etc., are not the only possible representatives for a simple notion of growth. For example, a production index could be used. All, however, would have to be corrected for price changes, a particularly important adjustment since "growth" implies necessarily long periods and there have been none in which there were no substantial price changes. A volume production index would have to be corrected for quality changes. Similarly, population growth has to be taken into account. Omission of this correction would be particularly serious since in some countries population has a tendency to outstrip the increase in production. Consider an extreme case where real gross national product increases by 10 percent per annum over a 3-year period, but population increases by 15 percent. Then the gross national product per capita would actually have decreased by about 5 percent per annum. This is a real possibility in populous underdeveloped countries, where it is notoriously difficult to obtain reliable data for production as well as for population. The possibility arises in particular with respect to Communist China, where

no one can say with confidence what the population is and how production has developed. From all we know, the likelihood of a net decrease in income per capita cannot be ruled out. Note that the question of optimal or just distribution of the national product is usually evaded in these discussions. The matter is obviously controversial and cannot be settled before a true theory of social benefit with a method of its measurement has been established.[1] Similarly, the age distribution of the population should not be neglected in evaluating size, change, and meaning in real national product.

We are here only concerned with the increases of real gross national product, i.e., the original gross national product series published by the Department of Commerce and deflated by various price indexes calculated by the National Income Division. These deflators are especially constructed price indexes for the various components of gross national product.

3. The Accuracy of Growth Rates.

Obviously, the value of a growth rate depends on both the accuracy of the figures for gross national product and of the prices going into the construction of the deflator-indexes. The former are subject to the considerable uncertainties discussed in the preceding section, the latter depend on the precision with which actual prices as distinguished from posted prices, list prices, etc., can be determined and applied to the correct sectors of gross national product. We know that this is far more difficult than generally assumed, but we shall discuss neither basic price data nor price indexes any further. The two factors do, however, combine in producing very serious doubts about the reliability and usefulness (in the current sense) of growth rates.

---

[1] This was already commented upon in the preceding section. We are, of course, far from this state.

The reader who concurs in the preceding evaluation of national income statistics will have no difficulty in seeing that a reliable growth rate of two significant digits is impossible to establish. But even the first digit is in grave doubt. This will be clearly shown below. Yet the emphasis of the public discussion is on the second digit — usually the first decimal — and it is carried on in all seriousness as if a distinction of, say, 3.2 and 3.3 percent were really possible, and as if the transition, within a short time, from the former to the latter constituted progress of the country, offered assurance of progress in the international competition, and so on. Such contentions are entirely unwarranted. It is difficult to see how even a shade of proof can be offered by the proponents of these practices. A growth rate simply cannot be computed with the stated or demanded degree of refinement and reliability. This applies to the existing national income data of any country in the world.

Yet we know that countries have grown and that, at periods, some have grown faster than others. But such observations and statements can be made with confidence only qualitatively and for longer periods. They are impossible to make for one year (or less!), where a nation's growth is as imperceptible as the growth of a person's teeth in a month. These general statements are based on a host of qualitative and quantitative indications, of which the imperfectly measured change in gross national product or national income is only one. Many are essentially qualitative in nature for which, as yet, no measurements have been devised, such as the development of institutions of business, markets, law, enterprise, etc. By denying the alleged accuracy of precisely stated growth rates we apply a standard that has been used throughout this investigation. In addition we express great doubt that as complex a phenomenon as "growth" can be stated adequately by as simple and almost trivial a measure as a percentage change in

either gross national product or national income, even if reduced to real terms. This doubt, I am sure, must be shared in their hearts by numerous theorists who have tried to explore in depth the intricate phenomenon of economic development, growth, and expansion. Their work has led to some non-trivial models which for their empirical application would require statistical measurements that differ by a wide margin from the simple percentage figures of a deflated gross national product.

Concerning the rates themselves, we observe the following: Table G-1 shows growth rates as commonly computed, but for 1, 3, and 5 percent plus or minus variations of the underlying figures. We recall that the assumption of a $\pm$ 5 percent accuracy of the non-deflated gross national product is conservative. The results of this simple computation should shake the confidence of anyone who thinks that the difference between, say, 3.2 and 3.3 percent is significant.

The computation is for a (hypothetical) change in United States gross national product from $550 billion in Period I to $560 billion in Period II. The first column lists the values of gross national product assuming the reported figure for Period I, i.e., $550 billion, to be subjected to the above-mentioned error of $\pm$ 1%, $\pm$ 3%, and $\pm$ 5% . The top row carried the same assumption through for the Period II figures. The body of the table contains the growth rates obtained for all combinations of the assumed possible errors. When there is no error assumed or when a positive error of a given magnitude is exactly compensated by an error of the same magnitude but with opposite sign, the growth rate is 1.8%. This rate would, according to current practices, be reported (and analyzed!) as "the" rate. It is, of course, impossible that there be no errors at all, and most improbable that they always exactly compensate for each other. The table now shows clearly what happens when even the modest 1% or 3% errors

Table G-1

Apparent Rate of Growth for ± 1, ± 3, ± 5 Percent Errors

A. Assuming Reported Gross National Product Figures 550 and 560 in Two Successive Periods

Period I GNP 550 ± Error	Period II GNP 560 ± Error	532.0	543.2	554.4	560.0	565.6	576.8	588.0
	% Error	- 5	- 3	- 1	0	+ 1	+ 3	+ 5
522.5	- 5	1.8	4.0	6.1	7.2	8.2	10.4	12.5
533.5	- 3	- .3	1.8	3.9	5.0	6.0	8.1	10.2
544.5	- 1	-2.3	- .2	1.8	2.9	3.9	5.9	8.0
550.0	0	-3.3	-1.2	.8	1.8	2.9	4.9	7.0
555.5	+ 1	-4.2	-2.2	- .2	.8	1.8	3.8	5.9
566.5	+ 3	-6.1	-4.1	-2.1	-1.2	- .2	1.8	3.8
577.5	+ 5	-7.9	-5.9	-4.0	-3.0	-2.1	- .1	1.8

Computed rate of growth assuming the reported figures to be correct is 560/550 = 1.8% .

B. Assuming Reported Gross National Product Figures 550 and 566.5 in Two Successive Periods

Period I GNP 550 ± Error	Period II GNP 566.5 ± Error	538.2	549.5	560.8	566.5	572.2	583.5	594.8
	% Error	- 5	- 3	- 1	0	+ 1	+ 3	+ 5
522.5	- 5	3.0	5.2	7.3	8.4	9.5	11.7	13.8
533.5	- 3	.9	3.0	5.1	6.2	7.3	9.4	11.5
544.5	- 1	-1.2	.9	3.0	4.1	5.1	7.2	9.2
550.0	0	-2.2	- .1	2.0	3.0	4.1	6.1	8.2
555.5	+ 1	-3.1	-1.1	1.0	2.0	3.0	5.0	7.1
566.5	+ 3	-5.0	-3.0	-1.0	0.0	1.0	3.0	5.0
577.5	+ 5	-6.8	-4.8	-2.9	-1.9	- .9	1.0	3.0

Computed rate of growth assuming the reported figures to be correct is 566.5/550 = 3.0% .

are introduced. Magnitudes and even signs are affected. If we assume that the reported figure of 550 for Period I is 5% too high and the figure for Period II 5% too low, we arrive, instead of at 1.8%, at 12.5% as the growth rate. If we reverse the assumption, the growth rate is - 7.9%. Suppose gross national product for the second year is only + 1% off and gross national product for the preceding one - 1% (a total error of only 2%), then the growth rate is 3.9%, but if the signs of the errors are reversed, the growth rate is - 0.2%! It is in the essence of an error estimate that a positive and negative deviation has to be admitted. Surely, the assumption of only a 1% error for each period is a very mild one. (Recall that the best British estimate is up to ± 3% and the best average Kuznets estimate is 7.5%!) The reader should contemplate what this trifling difference in our assumption entails. If our basic figures of 550 and 560 are more than 1.8% apart, say 3%, the results of a corresponding table are necessarily worse. For example, a - 1% error in the first period and a + 1% error in the second then give a growth rate of 5.1%, and if the signs are reversed a growth rate of 1.0%. With ± 3% the corresponding figures are 9.4% and - 3.0% respectively.

It is easily shown that the computations of Table G-1 are independent of the absolute amount of the assumed level of gross national product, and that the rates depend solely on the percentage change of gross national product from Period I to Period II and the errors. Moreover, the computations obviously apply to any situation where rates of change are involved and where the data are subject to error. In other words they apply to all economic data.

This simple arithmetical exercise combined with the indisputable fact that our final gross national product or national income data cannot possibly be free of error[1] raises the question whether the computation of

[1] Nor can it be assumed without further proof that the errors remain constant over time, that they change uniformly over time, and that the signs of the error never reverse themselves.

growth rates has any value whatsoever.

The following Tables G-2 show the results of estimating the <u>compound annual rate of growth</u>, r, on the basis of two observations, I_0 (the initial observation) and F_0 (the final observation) after n years, with percentage errors of E_I and E_F, respectively, where

$$I_0 = (1 + \frac{E_I}{100}) \times \text{(actual figure initially given)},$$

$$I_F = (1 + \frac{E_F}{100}) \times \text{(actual figure after } n \text{ years)}.$$

We may use as the estimated rate of growth:

$$r_E = 100 \left[\sqrt[n]{\frac{(1 + \frac{E_F}{100})(1 + \frac{r}{100})}{(1 + \frac{E_I}{100})}} - 1 \right] ,$$

which may be approximated by

$$r_E \cong r + \frac{E_F - E_I}{n} .$$

It should be noted that the estimate r_E becomes closer to r as n gets larger but loses its realistic significance simultaneously. Also, for fixed errors E_F, E_I the part of r_E contributed by these errors gets smaller proportionately, as r gets large.

The apparent annual rate of growth is given for $r = 1.8$, $r = 3.0$, and for the various values of E_F and E_I, in the cases $n = 5$ and $n = 10$.

Table G-2

Apparent Compound Rate of Growth in Percentage

A. Where n = 5 (r = 1.8)

E_F =	E_I = - 5%	- 3%	- 1%	0%	+ 1%	+ 3%	+ 5%
- 5%	1.8	1.4	1.0	.8	.6	.2	- .2
- 3%	2.2	1.8	1.4	1.2	1.0	.6	.2
- 1%	2.7	2.2	1.8	1.6	1.4	1.0	.6
0%	2.9	2.4	2.0	1.8	1.6	1.2	.8
+ 1%	3.1	2.6	2.2	2.0	1.8	1.4	1.0
+ 3%	3.5	3.0	2.6	2.4	2.2	1.8	1.4
+ 5%	3.9	3.4	3.0	2.8	2.6	2.2	1.8

B. Where n = 10 (r = 1.8)

E_F =	E_I = - 5%	- 3%	- 1%	0%	+ 1%	+ 3%	+ 5%
- 5%	1.8	1.6	1.4	1.3	1.2	1.0	.8
- 3%	2.0	1.8	1.6	1.5	1.4	1.2	1.0
- 1%	2.2	2.0	1.8	1.7	1.6	1.4	1.2
0%	2.3	2.1	1.9	1.8	1.7	1.5	1.3
+ 1%	2.4	2.2	2.0	1.9	1.8	1.6	1.4
+ 3%	2.6	2.4	2.2	2.1	2.0	1.8	1.6
+ 5%	2.8	2.6	2.4	2.3	2.2	2.0	1.8

Table G-2

Apparent Compound Rate of Growth in Percentage

C. Where n = 5 (r = 3.0)

E_F =	E_I = - 5%	- 3%	- 1%	0%	+ 1%	+ 3%	+ 5%
- 5%	3.0	2.6	2.2	2.0	1.8	1.4	1.0
- 3%	3.4	3.0	2.6	2.4	2.2	1.8	1.4
- 1%	3.9	3.4	3.0	2.8	2.6	2.2	1.8
0%	4.1	3.6	3.2	3.0	2.8	2.4	2.0
+ 1%	4.3	3.8	3.4	3.2	3.0	2.6	2.2
+ 3%	4.7	4.2	3.8	3.6	3.4	3.0	2.6
+ 5%	5.1	4.7	4.2	4.0	3.8	3.4	3.0

D. Where n = 10 (r = 3.0)

E_F =	E_I = - 5%	- 3%	- 1%	0%	+ 1%	+ 3%	+ 5%
- 5%	3.0	2.8	2.6	2.5	2.4	2.2	2.0
- 3%	3.2	3.0	2.8	2.7	2.6	2.4	2.2
- 1%	3.4	3.2	3.0	2.9	2.8	2.6	2.4
0%	3.5	3.3	3.1	3.0	2.9	2.7	2.5
+ 1%	3.6	3.4	3.2	3.1	3.0	2.8	2.6
+ 3%	3.8	3.6	3.4	3.3	3.2	3.0	2.8
+ 5%	4.0	3.8	3.6	3.5	3.4	3.2	3.0

It is clear from the above that the usefulness of growth rates is not increased when compound rates are considered. We continue, however, by looking back at Table N-2, which showed the revisions throughout the years of initially given estimates of national income. Here the question of the growth rate is raised with respect to the particular moment in time when it is computed. Table G-3 shows the results. If the rate is determined for the change from 1947 to 1948 in February 1949, when the first figures became available, it was 10.8%. In July 1950, using officially corrected figures, it became 12.5%; in July 1956 it fell to 11.8%, only to rise again to 12.8% in July 1958 — a full percentage point! — the highest of all values. All this for the growth from 1947 to 1948! Similar observations apply to the other years for which this computation has been made. There is no consistency in the changes. In stating what the growth rate of the country is, much depends, therefore, on the moment of time when a growth rate is computed. Though not surprising in the light of our previous investigations, this result is nevertheless noteworthy. And all this applies to figures where we have abstracted from the fact that they are necessarily afflicted with errors which, when low, must be at least 5%. If we make allowance for errors, however slight, the confusion mounts.

To round off this picture we refer to Table G-4, where the growth rates from year to year are computed by taking alternatingly (from Table N-2) the lowest and highest estimate of national income for one year and determining the growth rate to the highest and lowest estimate of the following year. Such differences exist and are sometimes appreciable, running in absolute terms up to about $9 billion for one year. Clearly the lowest to highest ratio will be the largest of all four rates for the same pair of years, as the highest to lowest will give the smallest rates, with the others in between. The differences are enormous (for the same pair of

Table G-3

United States Growth Rates

(Based on Consecutive Corrections of Original Estimates of National Income)

	1947-1948	1948-1949	1949-1950	1950-1951	1951-1952	1952-1953	1953-1954	1954-1955	1955-1956	1956-1957	1957-1958	1958-1959	1959-1960
Feb. 1949	10.8												
July '50	*12.5	-3.0											
July '51	12.5	-3.0	10.3										
July '52	12.5	*-3.2	*10.6	16.0									
July '53	12.5	-3.2	*11.2	*15.7	4.7								
July '56	*11.8	*-2.4	*11.0	*15.4	*4.8	4.1	-1.3	8.6					
July '57	11.8	-2.4	11.0	15.4	4.8	4.1	*-1.0	*8.4	6.0				
July '58	*12.8	*-2.6	*11.1	*15.5	*4.6	*4.6	*-1.2	*9.4	*5.8	4.2			
July '59	12.8	-2.6	11.1	15.5	4.6	4.6	-1.2	9.4	*6.2	*4.5	-0.1		
July '60	12.8	-2.6	11.1	15.5	4.6	4.6	-1.2	9.4	6.2	*4.6	*0.2	*8.7	
July '61	12.8	-2.6	11.1	15.5	4.6	4.6	-1.2	9.4	6.2	4.6	*0.1	*8.8	4.4

Note: Asterisk indicates change from previous value.

Source: Table N-2.

Table G-4

United States Growth Rates

Computational Effect of Utilizing High/Low Year-to-Year Estimates

1947-1948	1948-1949	1949-1950	1950-1951	1951-1952	1952-1953	1953-1954	1954-1955	1955-1956	1956-1957	1957-1958	1958-1959	1959-1960	1960-1961
						Lowest to Lowest							
12.4	-2.4	9.0	17.1	5.0	4.4	-1.3	3.6	5.7	4.6	0.7	10.5	4.7	3.1
						Highest to Highest							
11.6	-2.1	9.2	15.5	4.6	5.3	-1.9	9.4	6.2	4.6	0.2	8.7	4.4	3.1
						Highest to Lowest							
9.4	-4.4	6.4	14.0	3.7	3.4	-2.6	7.4	3.7	2.1	-1.7	8.4	4.4	3.1
						Lowest tc Highest							
14.7	0.0*	11.9	18.5	5.9	6.3	-0.1	10.7	8.3	7.2	2.7	10.8	4.7	3.1

*Not significant.

years), sometimes even the sign is switched! There is an apparent decrease of the difference for the last two years (1959 to 1960 and 1960 to 1961). But we do not know, at this time of writing, what surprises the revisions of these figures will bring during the next 8-10 years, if the experience of the past years, as seen from Table N-2, is a guide for our expectations. Again, it has to be mentioned that this computation does not take into consideration errors; if that were done, as it should be, the "growth rates" would jump around even more wildly.

The computations of Table G-1 should be applied, mutatis mutandis, to all other countries. Identical considerations are therefore valid regarding the growth rates of other nations, where the presumed average error will often exceed 5% by a wide margin. This was true even of Great Britain. One can well imagine what the value of growth rates of statistically less developed countries must be like.

4. The Choice of the Base Year.

In addition to all these difficulties there is the ambiguity in choosing the base year. The need for a base year arises from the desire to compare long periods by means of the compound rate. Such periods will often comprise a series of business cycles and therefore several decades. If a year with a high (low) gross national product is chosen as base year, this will depress (raise) the growth rate of subsequent years. Since there is no such thing as a "normal" year, the investigator has a great amount of freedom in determining a base year. An unscrupulous or politically oriented writer will choose that base year which produces the sequence of (alleged) growth rates best suited to his aims and programs. An advocate of government spending and inflation will pick a year with a high gross national product as base year in order to show a low rate of growth and thereby to

strengthen his argument in favor of inflation, government deficits, and the like. An opponent of such policies will choose a relatively poor gross national product as base year, thus obtaining a series of growth rates carrying the comforting message that the development of the country is progressing well. These are, of course, standard tricks, used, undoubtedly, ever since index numbers were invented.

Consider the following:

Table G-5

United States Growth Rates Compounded from Different Base Years
(Gross National Product in 1954 Dollars)

Growth from	Rate
1949 to 1960	3.7 per annum
1950 " "	3.3 " "
1954 " "	3.2 " "
1955 " "	2.3 " "

(In interpreting this table we make — for the sake of argument — the assumption that the growth rates are significant even to two digits; it was shown above that this need not be so.) Suppose a 3.5% growth is considered desirable: then only 1949 chosen as a base year shows that the goal has been reached; if others are chosen, a failure has to be recorded. 1949 and 1954 were recession years, in 1950 occurred the outbreak of the Korean War, and 1955 was a year of high business activity. We dispense with any more elaborate illustrations.

We have limited ourselves to growth rates based on gross national product, etc., but our remarks apply also to any of the many other widely used rates. To name only a few, there is investment as a percentage of national income, the value added by manufacturing as a percentage of total

value added, exports plus imports as a percentage of national income. There are countless others. In fact, the above considerations apply no matter what the substratum of computed rates of change. They may be data on production, foreign trade, prices, turn-overs, etc., in short anything the economist has to deal with. All are based on empirical data and all have some error component. If the rates and the changes in the rates could be obtained in a reliable manner, economic analysis and economic policy would benefit immeasurably. They are, unfortunately, being computed freely[1] and are used indiscriminately as if no problem existed regarding their accuracy and therefore their value.

In addition to these indiscriminate uses, as expected, almost any argument can be supported by inobtrusive manipulation. The literature is full of examples, and although the matter is trivial, these abuses will continue as long as there are unscrupulous investigators and gullible readers.

5. International and Inter-Period Comparisons of Growth Rates.

The difficulties discussed in § 9. of the preceding section, relating to comparisons of national incomes for different countries and different periods of time even for a single country, reappear when growth rates are considered. There are, essentially, two observations, one pertaining to the data, the other to the concepts.

First, it is clear, and now requires no further comment, that in view of the high degree of unreliability of basic national income data the growth rates even for the United States are at best very shaky. From their relatively better wuality we have to go over to the lesser and lesser quality of the national income statistics of other countries — even the United Kingdom would rank lower — until the 38 "fair and weak" countries

[1] Cf. Yearbook of National Account Statistics, United Nations.

of Table N-8 are reached. The computation, and hence the comparison, of international growth rates under these conditions is a most dubious undertaking. Even gross differences do not reliably point to the true character of the underlying processes. Perhaps when very gross differences persist between two countries, say one showing about 10 percent growth and the other 2 percent, and if the general level of statistical reporting for both is approximately at the same high level, can one say that the former's gross national product has risen faster than the latter's, but when the reported rates are very close to each other hardly any conclusions can be supported as scientifically acceptable.

It will always be necessary to supplement the rates by qualitative information as indicated above. We emphasize again: there can be no doubt that countries develop at different speeds and that this fact is noticeable over longer periods of time, particularly when the initial level of economic activity is low and the state of technology is primitive. When big gaps exist, a comparison of change can be made with some confidence provided a sufficiently long time interval is admitted. When countries are very similar in their structure, such comparisons become immeasurably more difficult and unreliable. It is ironical that the differences among highly developed and similar countries are harder to ascertain than for less advanced nations, especially when the latter's growth is compared to that of the former.

Second, the last remark points up the conceptual difficulties in comparing growth rates — assuming that they could be computed in the current manner from gross national product. It is doubtful that the same ratio computed for very different kinds of countries is equally meaningful. If there is any value at all to the notion that countries grow in characteristic patterns, depending on their history, technological age, geographic

position, size, etc., then it is unlikely that a single simple number can state adequately, or at least not in a misleading sense, how they evolve relatively to each other. In this respect the problems of finding a proper solution for describing the gross national product or the national income or developing an even more suitable concept for a given moment in time are compounded many times. It is well known that the problem of finding a social benefit or welfare function so far is unsolved. Hence the hope that we could make meaningful international comparisons over time is vain when attempted in terms of such simple and unreliable percentages.

A particular fallacy in using growth rates for comparing different countries needs to be explained: The growth rates are frequently used as measures for determining the variations in strength or power of different countries in world affairs or in respect to the cold war situation. There is a conceptual problem involved that would have to be taken care of even if the statistics were in good shape. It is clearly possible to have two countries with the <u>same</u> growth rate, but where country A expands by adding to its output of automobiles, refrigerators, swimming pools, etc., while country B increases its output of machine tools, power plants, mines, etc. The second is laying the foundation for further output increase while the first is not. Similar considerations apply when weapons and other tools for war are involved. The ordinary growth rate, computed for the big gross national product aggregate, covers up these profoundly different developments and would easily give entirely erroneous and misleading information about the relative development of these countries. Yet this is the figure commonly used to assess past progress and future tendencies. The answer would be to compute instead "power indexes" (of growth) which would have to be based on the information given by special aggregates made up of better related

components.[1]

To sum up: There is no possibility of making concessions as far as the scientific use of growth rates is concerned. As available today, they are worthless in view of the exacting uses to which they are being put. The data are limited and untrustworthy, and the method of computation is at best based on the tremendous oversimplification that there are no errors in gross national product.

To put this differently: Precise uses of "growth rates" are entirely inadmissible, whether for comparing different countries or short periods of the same country. Their computation is largely arbitrary. The concept itself is vague and unreliable. Anyone using growth rates in the current manner will have to show that the above arguments are all taken into consideration and the corresponding objections have been overcome.

A recent example of the extravagant uses of growth rates is found in the Proceedings of the American Economic Association (Vol. LII, May 1962) reporting on the session on "The Lagging U. S. Growth Rate."[2] There no reference whatsoever is made to the accuracy of the underlying data and to the reliability of the various growth rates discussed. The chief speaker's

---

[1] Cf. O. Morgenstern: The Question of National Defense, 1959, Second Revised Edition, Vintage Books V-192, New York, 1961, pp. 202 ff.

[2] E. F. Denison, "How to Raise the High-Employment Growth Rate by One Percentage Point," loc. cit., pp. 67-75. The author is with the Committee for Economic Development, a generally valuable policy oriented group, supposedly concerned with the realities of the American Economy. In his discussion, G. Colm states (p. 87) that Denison's paper is "statistically well supported." Clearly this is in complete contradiction to our findings. In another publication Mr. Denison does recognize several of the problems connected with growth rates raised in this chapter, such as the effect of the choice of period over which the rate is computed, the large effects of even small errors in the basic data, and the consequent meaninglessness of discussing fairly small changes in the rate of growth, especially over periods which are not long. Cf. Edward F. Denison, The Sources of Economic Growth in the United States and the Alternatives Before Us, Supplementary Paper No. 13, Committee for Economic Development (Washington, 1962), pp. 16-19.

assignment, given him by the President of the Association, was "to devise a package of proposals that can raise the growth rate over the next twenty years by one percentage point" (p. 67). He then assumes a rate of $3\frac{1}{3}$ % to be raised to $4\frac{1}{3}$ % . This leads to a search of components for a 13-part program in which the contribution of <u>each</u> proposal "will be stated in <u>hundreds</u> (!) of a percentage point" (p. 70). This is clearly impossible both as far as goals and means are concerned. These authors obviously have before their minds a picture of the American (or any other) economy given with a precision and detail that has only a remote resemblance to the true picture we are capable of drawing and is revealed by even a moderately critical examination of the data.

We conclude that growth rates as commonly computed from gross national product and national income data whose errors are known to be large, though they are not stated numerically by their makers, are completely worthless as far as international comparisons are concerned. International comparisons of the relative growth and development of different countries demand, in addition to qualitative, historical, sociological description, more specific indexes, carefully constructed from those activities that are relevant to these countries in regard to their location, climate, technology, policies, etc. Such indexes would, of course, also exhibit some of the faults described above, but they could perhaps be brought under some measure of control.

It would, indeed, be most peculiar if the economic "growth" of nations could be described with virtually no effort by merely comparing consecutive numbers, these being their gross national product or something similarly comprehensive. Should economics ever reach the pleasant state where "growth" can be measured reliably by a <u>single</u> (!) number, this condition will only be achieved after developments in this science have taken

place which now cannot be foreseen.

It remains to point to the relation of the above results to the various theoretical models of economic growth. Their value lies, of course, in their conceptual construction. But they have all been devised with an expectation of application. They all contain rates of one kind or another and all make statements about stability, equilibria, and the like. It will be a long way before they can be applied — whatever their merits otherwise — should the data they require exhibit characteristics of the kind described in these pages.

FOR ERRATA

AD 285 065

THE FOLLOWING PAGES ARE CHANGES

TO BASIC DOCUMENT

Errata Sheet for Research Memorandum 43

in section: "The Validity of Economic Growth Statistics and Growth Rates"

Page 8, line 2 - change "high" to "low"

Page 8, line 3 - change "low" to "high"

Page 9, the second factor of the numerator of the fraction under the nth root sign, used in the estimation of the rate of growth should read:

$$(1 + \frac{r}{100})^n$$

rather than:

$$(1 + \frac{r}{100})$$

END CHANGE PAGES